Net-zero Building Cluster Simulations and On-line Energy Forecasting for Adaptive and Real-Time Control and Decisions

A Thesis

Submitted to the Faculty

Of

Drexel University

by

Xiwang Li

in partial fulfillment of the degree

of

Doctor of Philosophy

Oct. 2015

For Dissertation Approval From

DEDICATION

To my parents and loved ones.

ACKNOWLEDGEMENTS

The support from many individuals made this thesis happened. First of all, I would like to thank my academic advisor, Dr. Jin Wen, for granting me the opportunity to pursue the doctoral degree and her continuing mentoring, guidance and support. Over the past five years, she taught me how to conduct research and how to become a researcher. Her professional insight, open-mind research expertise, and caring personality helped and lead me through my PhD study. I look forward to the future success of the Building Science & Engineering Group (BSEG) under her capable leadership.

My special thanks also go to my Ph.D. committee members Dr. Er-Wei Bai (University of Iowa), Dr. Patrick Gurian (Drexel University), Dr. Micheal Waring (Drexel University), and Dr. Teresa Wu (Arizona State University). They provided valuable recommendations and comments for my Ph.D. research and my dissertation. Their diverse sets of expertise added depth to the dissertation. Besides my committee members, I also worked with Dr. Kemper Lewis (University at Buffalo-SUNNY) during the past three years on my thesis project. He provided meaningful insights and improved the quality of my study.

The findings in this thesis are validated and verified at the Iowa Energy Center. I would like to thank Dr. Xiaohui Zhou and Dr. Ran Liu from Iowa Energy Center for their help and guidance during my real field experiment verification period. During my internship at Siemens Corporation, I was supervised by Dr. Zhen Song and Dr. Sanjeev Srivastava. I feel a deep gratitude to them for providing me the opportunity and their supervision during my internship.

I would like to express my thanks to my collaborators: Cui Can (Arizona State University) and Philip Odonkor (University at Buffalo-SUNNY). I had a very pleasant working experience with them. I also would like to thank all my BSEG lab mates and friends at Drexel: Liam

Hendricken, Jared Langevin, Qiang Mao, Shokouh Poorarian, Adam Regnier, Adams Rackes, Yuxing Sun, Sheng Wang, Chunyi Wang, and Yanan Yang for the great friendship and collaboration during my time at Drexel. I wish all of them the best of luck in their future careers.

I cannot adequately express my gratitude to my parents, far away in China, for their unconditional support and indescribable sacrifices that they made in affording me to pursue my dreams. I wish that I have made them proud.

Finally, I would like to acknowledge the National Institute of Standards and Technology (ID: 60NANB10D271) and National Science Foundation (ID: 1239247) for their finical support throughout my Ph.D. study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Net-zero Building Cluster Simulations and On-line Energy Forecasting for Adaptive and Real-Time Control and Decisions

Xiwang Li

Jin Wen, Advisor, PhD

Buildings consume about 41.1% of primary energy and 74% of the electricity in the U.S. Moreover, it is estimated by the National Energy Technology Laboratory that more than 1/4 of the 713 GW of U.S. electricity demand in 2010 could be dispatchable if only buildings could respond to that dispatch through advanced building energy control and operation strategies and smart grid infrastructure. In this study, it is envisioned that neighboring buildings will have the tendency to form a cluster, an open cyber-physical system to exploit the economic opportunities provided by a smart grid, distributed power generation, and storage devices. Through optimized demand management, these building clusters will then reduce overall primary energy consumption and peak time electricity consumption, and be more resilient to power disruptions. Therefore, this project seeks to develop a Net-zero building cluster simulation testbed and high fidelity energy forecasting models for adaptive and real-time control and decision making strategy development that can be used in a Net-zero building cluster.

The following research activities are summarized in this thesis: 1) Development of a building cluster emulator for building cluster control and operation strategy assessment. 2) Development of a novel building energy forecasting methodology using active system identification and data fusion techniques. In this methodology, a systematic approach for building energy system characteristic evaluation, system excitation and model adaptation is included. The developed methodology is compared with other literature-reported building energy forecasting methods; 3) Development of the high fidelity on-line building cluster energy forecasting models, which

includes energy forecasting models for buildings, PV panels, batteries and ice tank thermal storage systems 4) Small scale real building validation study to verify the performance of the developed building energy forecasting methodology. The outcomes of this thesis can be used for building cluster energy forecasting model development and model based control and operation optimization. The thesis concludes with a summary of the key outcomes of this research, as well as a list of recommendations for future work.

1. CHAPTER 1 Introduction

1.1 Background

The electricity consumption of the US grew 1.7% annually from 1996 to 2006, and the total growth will reach 26% until 2030 [1]. Among that consumption, buildings are responsible for over 70% of electricity consumption in the US [2]. Around 30% of the energy used in building is consumed by heating, ventilating and air conditioning (HVAC) [3]. As a results, building energy efficiency and Net-Zero building have become international targets for building design and operation since last decade. The U.S. Energy Independence and Security Act of 2007 specified that the net zero energy commercial building initiative to support the goal of "Net-Zero for all new commercial buildings by 2030", "Net-Zero for 50% of U.S. commercial buildings by 2040", and "Net-Zero for all U.S. commercial building by 2050 [4] . The European Union also establishes the 'nearly Net-Zero energy building' as the building target from 2018 for all public owned or occupied by public authorities buildings and from 2020 for all new buildings [5]. Building design and operation are the two essential factors for the realization of Net-Zero buildings. Building insulation, shading, and equipment efficiency are fundamentals of energy saving and net zero energy in building design perspective. The operation of the building HVAC system, on-site energy generation and storage system are the opportunities for net zero energy in building operation [6]. Moreover, it is estimated by the National Energy Technology Laboratory that more than one-fourth of the 713 GW of U.S. electricity demand in 2010 could be dispatchable if only buildings could respond to that dispatch through advanced building energy control and operation strategies and smart grid infrastructure [7]. The building operation is therefore significant in terms of decreasing the energy consumption and peak demand. Currently, with the development of smart grid, the power infrastructure in U.S. is experiencing a

revolutionary transformation, from a centralized one-way communication to a decentralized network with two-way communication. With the trend of moving from centralized building operation decision to decentralized operation control, it is envisioned that neighboring buildings will have the tendency to form a building cluster, within which smart grids, distributed power generation, and storage devices, can freely share energy resource locally and globally and the entire cluster will achieve maximum energy efficiency. It is anticipated that this building cluster concept will fundamentally transform the energy industry by shifting expensive on-site energy generation aimed at creating single Net-Zero building one-at-a-time to an autonomous and adaptive system of buildings aimed at Net-Zero clusters. Currently, as more and more distributed energy generation and storage system penetrate the power grids and different micro-girds, such as building cluster networks, appear, the energy (cost) saving potentials from the interaction of different systems is increasingly progressing. Transitive energy, known as "*A set of economic and control mechanisms that allows the dynamic balance of supply and demand across the entire electrical infrastructure using value as a key operational parameter",* is now starting to change the grid and building operation. The emerging and improvement of smart grids, Net-Zero buildings, and especially demand response technologies drive a revolutionary transformation of the power infrastructure, from a centralized network with one-way control to a decentralized network with two-way interactive communication and control. This transformation provides an opportunity to explore the energy (cost) saving potentials of Net-Zero building clusters. It is envisioned that similar to an ecological system, when individual buildings are allowed to freely exchange energy and information, the entire power grid will converge to a state of maximum efficiency.

To realize the transactive energy and building cluster vision, high fidelity building energy model is the most critical component. To develop operation strategies for transactive energy and

building cluster vision, simulation testbed that can simulate the building cluster, provide real-world-like operation data, and assess different operation strategies is also essential. Therefore, the research of this project will follow these two directions. Extensive studies exist in the literature on building energy simulation and forecasting. A detailed literature review will be presented in next section.

1.2 Literature Review

Over the past decades, two groups of building energy models have been developed for different purposes: building design and control (operation). These models can be categorized in many different ways. Based on their applications, they can be categorized as design models and operation models, based on their targeting systems, they can be categorized as whole building models and component models, based on their usage, they can be categorized as offline and online models, and based on their calculation theories, they can also be categorized as white, black and gray box models [8].

Models for design are usually used to determine the building load, equipment size, and to evaluate the energy input to assist system design, while models for operation simulate the behaviors of the building energy systems for the control and operation systems to save energy and improve indoor comfort. From the modeling targets, the models can be categorized into whole building models and building component models. Whole building model focuses the overall energy simulation and forecasting for a building as a whole, while building component model simulates the behavior of a specific building component, such as AHU, PV panel, and battery, etc.. From different application scope, the models can be categorized into global energy management models and local control models. Usually, models for global energy management are the models for whole building simulation. They need to forecast the operation of the whole

building to manage the building overall operation. The models for local control are component models, which simulate the detailed operation of specific building devices for better control and operation. From the different model developing methods, they can be categorized into white box models, black box models and grey box models. White box model uses detailed physical equations to simulate the behaviors of building, building system, building equipment, and etc. It is more complicated and accurate than the other two, while it is more computation demanding. Black box model, also known as data driven model, uses statistical methods to capture the relationship between the input and output variables. Grey box model is a hybrid model of white box and black box model, which applies simplified physics theory to develop the model structure and uses data driven methods to identify the model parameters. All these types of models have their pros and cons, and their best application areas. The research scope of this project is to improve the operation of building cluster. Therefore, in this section, all three modeling approaches, namely, white box, grey box and black box approaches, are reviewed. Several topics related to high fidelity model development, i.e., using system identification method for model development, data fusing, and co-simulation environment, are also reviewed.

1.2.1 White Box Simulation Models

White box (or forward) modeling approach uses detailed physics based equations to model building components, sub-systems and systems to predict whole buildings and their sub-systems behaviors, such as their energy consumption and indoor comfort. Due to the detailed dynamic equations in white box models, they have the potential to capture the building dynamics well but they are time consuming to develop and solve. White box modeling approach can be used to model one component, a subsystem or a comprehensive system, such as a whole building, including both building envelope and its mechanical systems.

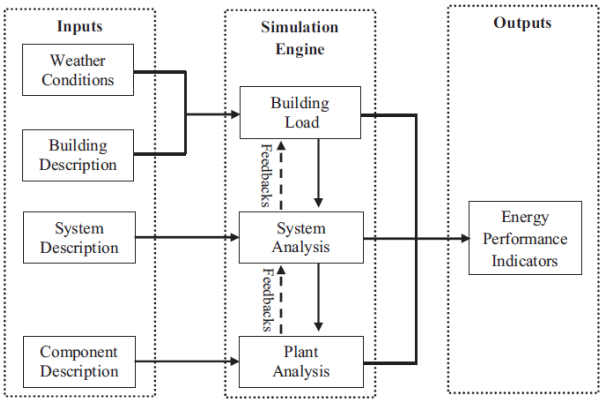


Figure 1-1. General data flow and main procedure of detailed simulation

In a whole building white box simulation process, the general data flow and main procedure of white-box model development and simulation are summarized in Figure 1-1. The parameters for weather condition, building structure, building systems and building equipment need to be obtained from their physical characteristics, usually from design plan, manufacture catalog or on-site measurement. The simulation engine is a group of mathematical equations which simulate the building operation and calculate the building energy consumption. There are many white box software tools that can be used for both whole building and component level simulation, such as EnergyPlus [9], TRNSYS [10], HVACSIM+ [11], Modelica [12], and eQuest [13]. These tools are firstly reviewed in this Section. The application of white box simulation tools in whole building and component simulation are reviewed following the tool introduction. Lastly, studies that apply white box models for building control and operation are discussed.

1.2.1.1 White Box Simulation Tools

EnergyPlus [9] is a whole building energy simulation program that engineers, architects, and researchers use to model energy and water use in buildings. Modeling the performance of a building with EnergyPlus enables building professionals to optimize the building design to use less energy and water. Thermal loads are calculated in EnergyPlus by a well-mixed heat balance

mode at user specified time steps and passed to a building systems manager, which provides feedback about the plant and electrical system responses and what portion of the loads (if any) have not been met. Any un-met portion of the loads would be reflected in the following time step through an adjusted space temperature. This incorporation of HVAC system feedback into EnergyPlus load calculations addressed a major deficiency in the BLAST and DOE-2 simulation programs. Figure 1-2 shows the structure of the EnergyPlus integrated solution manager that manages the surface and air heat balance modules and acts as an interface between the heat balance and the building systems simulation manager. Several graphical user interfaces have been added into different software for EnergyPlus, such as OpenStudio, SketchUp, for users to develop EnergyPlus model and analyze simulation data.

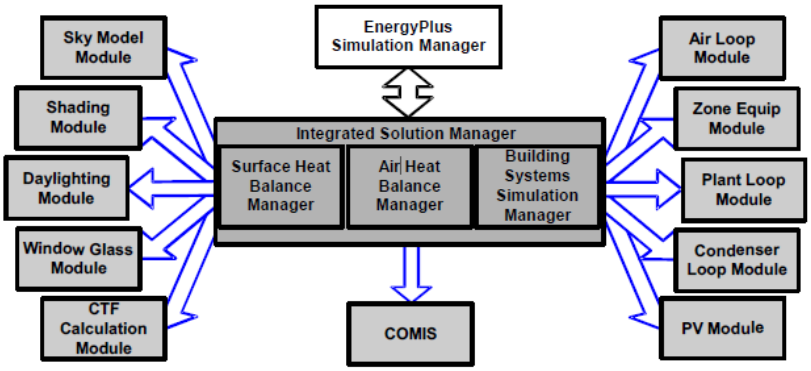


Figure 1-2. EnergyPlus simulation procedure

TRNSYS [10] is an flexible graphically based software environment used to simulate the behavior of transient systems. TRNSYS is made up of two parts. The first is an engine (called the kernel) that reads and processes the input file, iteratively solves the system, determines convergence, and plots system variables. The second part of TRNSYS is an extensive library of components, each of which models the performance of one part of the system. The standard library includes approximately 150 models ranging from pumps to multi-zone buildings, weather

data processors to economics routines, and basic HVAC equipment to cutting edge emerging technologies. Models are constructed in such a way that users can modify existing components or write their own, extending the capabilities of the environment (see Figure 1-3).

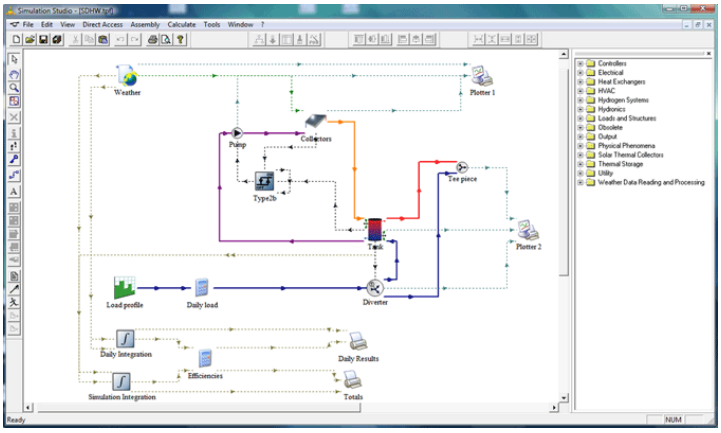


Figure 1-3. TRNSYS simulation interface [10]

HVACSIM+, similar to TRNSYS, is based on subroutines containing algorithmic models for the building systems' underlying physics. All the subroutines in HVACSIM+ are in three categories: preprocessing, simulation, and post-processing (see Figure 1-4). The main simulation program, MODSIM, is the equation solving engine. HVACGEN is the sub-model to generate the simulation file for the targeting building. The output of HVACGEN cannot be used by MODSIM directly, which need to be converted to definition file by SIMCON. The post-processing subroutine, SORTSB, can sort the output file of MODSIM by a user defined order for plotting. The weather conditions and other boundary settings are provided in Boundary Data File, which is inputted into MODSIM for simulation.

Modelica building library, developed by LBNL, is an open-source component library for building energy system using the equation-based object-oriented modeling language Modelica. This Modelica based building library is able to quickly add models of emerging technologies into

the simulation environment for performance assessment, to evaluate different promising alternative operation and control schemes, to interface with real facilities at the component and whole building level, which uses symbolic algebra tools to reduce the dimensionality of the coupled systems of equations that need to be solved for simultaneously. Because of the advantages listed above, Modelica building library is now started to be used in building model-based controls, fault detection and diagnostics.

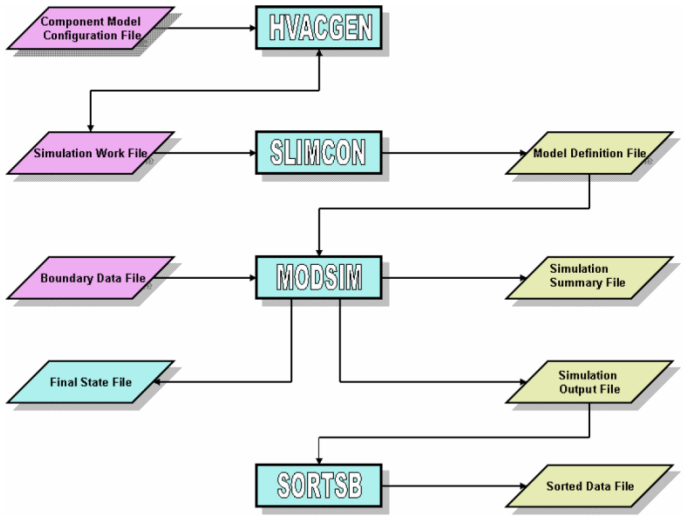


Figure 1-4. Flow chart of HVACSIM+ [11]

eQUEST [13] is an easy to use building energy simulation tool, which requires less expert work than EnergyPlus, while can provide professional-level simulation results. Therefore, it is more widely used in the real design and commissioning projects. eQUEST uses the same simulation engine as EnergyPlus, DOE-2, which can calculate a whole year heating or cooling loads, and also simulate the performance of other equipment, such as fans, pumps, chillers and etc.

There still are hundreds of white box simulation tools for building energy simulation, Crawley et al. [14] introduced and compared more than 20 different tools. However, in this project only the potential tools for this study are reviewed.

1.2.1.2 White Box Model for Whole Building Simulation

EnergyPlus, due to its high accuracy, has been widely used in building energy saving research. In order to provide a consistent baseline of building energy models, Department of Energy (DOE)'s building Technologies Program has developed a series of commercial building benchmark models [15]. These benchmark models are simulating different types of buildings, such as small office, large office, primary school, secondary school, shopping mall, and etc. EnergyPlus is a "text-based" simulation tool. The information of all the boundary conditions and building systems are created and stored in a "text-based" idf file. As a result, some "script-editing-based" optimization models have been developed to improve the building control and operation. The studies for building operation optimization using EnergyPlus will be reviewed in following section.

As a component based simulation tool, TRNSYS is more widely used in the building component simulation. Few studies have been found using TRNSYS to improve the whole building energy simulation [16-24]. Similar to TRNSYS, HVACSIM+ is also a component based simulation tool. However, it requires very high expert knowledge to develop and run the model, so it hasn't been commonly used in the existing studies. Several studies have been found using HVACSIM+ to simulate the building energy system. Shun and Wen [25] simulated and validated the VAV and fan coil system for a mid-size commercial building.

Modelica building library, due to its flexibility in creating building models and ability in accurately capture transients and state information, has been utilized in more and more studies for building energy simulation in the past 10 years [26-31]. Wetter et al. [26] modeled the heat transfer in buildings using Modelica building library. They discussed how to develop a single room model, how to compose a multi-zone building model, and how to decompose into sub-models. They also presented numerical experiments to validate this building library against the

selected cases of the ANSI/ASHRAE Standard 140-2007 envelop tests. Eisenhower et al. [27] developed and calibrated a control oriented dynamic model for a campus theater using Modelica. Two steps dynamic model calibration process were also introduced in this study. The first step was isolating steady state data and calibrating the flow network model. The second one is calibrating the parameters that pertain to the dynamic portions of the model. Based on Modelica building library, a hybrid model for energy saving study of subway stations was developed in [28]. This hybrid model combined CFD, lumped parameter model and stochastic modeling for the subway station temperature and energy consumption simulation.

eQUEST, due to its simplification in building energy models and calculation, it is widely used in practical building design, while rarely in building energy saving research. Few of them have been identified in this project [32-35]. As plug load energy use is claiming a larger percentage of the total building energy, Fuerters and Schiavon [34] studied the efffect of plug loads on the energy consumption in LEED certification and energy modeling. Heiple and Sailor [35] presented a framework combining building energy simulation with geographical information system. Prototype buildings were modeled using eQUEST, and the energy simulation data was matched with the existing buildings using geospatial mapping. Ke et al. [35] examined the impact of energy consumption parameter changes, such as occupancy, lighting power density schedule, equipment power schedule parameters, window types, and wall types, on the overall energy consumption for an office building in Taiwan. Using eQUEST, Yin et al. [36] conducted a case study for the energy saving potentials from solar window file in two commercial buildings in Shanghai. eQUEST was used to simulate the annual building performance with and without the solar window film. Measured monthly and daily electrical consumption were used to calibrate the simulation model.

1.2.1.3 White Box Model for Building Component Simulation

In this study, besides building envelope and typical building energy systems, alternative energy system, such as PV panel, and energy storage systems such as battery and ice tank, are used. Since these components/subsystems may not always be included in a whole building simulation tool, their models are specially reviewed on this section.

White box models for ice tank, PV panel and battery use detailed physics equations to represent the real behaviors of these systems, and then these equations will be solved by the white box tools' engines. EnergyPlus provides different models for ice tank, PV panel and battery based on different complicity. Similarly, TRNSYS also provides multiple models for PV panel and battery with different calculation theories. TRNSYS doesn't provide any default ice tank model, however, it allow users to create their own ice tank model according to their requirements. The detailed information and application of these white box component models will be discussed.

There are two different types of ice tank storage models in EnergyPlus, namely simple model, and detailed model. The detailed model allows user to define the charging and discharging characteristic curve, which will improve the simulation accuracy. Ihm et al. [37] developed an ice tank thermal storage module and integrated it with EnergyPlus, which is the default ice tank storage object in EnergyPlus now. In this study, they simulated base load operation of partial ice storage systems for a small building in Arizona. The results show higher chiller energy consumption for ice storage systems than that for non-storage systems for the design day. Henze et al. [38] developed and validated a simulation environment for the evaluation of the performance of various controls of ice storage system. Chiller-priority, constant-proportion, and storage-priority control strategies was compared to the optimal control strategy that achieves the theoretical maximum of operating cost savings. EnergyPlus ice tank storage model has also been used by Candanedo et al. [39] in developing a MPC for building energy cost saving. 5%-30%

energy cost saving was achieved in this MPC framework. Based on one of benchmark building EnergyPlus models provided in [15], Sehar et al. [40] analyzed the chiller energy consumption of conventional non-storage and ice storage cooling systems for large and medium-sized office buildings in different climate zones.

EnergyPlus provides three types of PV panel generator models: simple model, equivalent one-diode model, and sandia model. In Simple PV performance model, a constant efficiency assumed during whole range of solar irradiation and cell temperature effect has not been taken into account. Equivalent One-Diode model is known as four parameters model in which modules are modeled using an equivalent one-diode circuit. The Sandia model is incorporated which is based on empirical coefficients assembled by Sandia National Laboratory for each specific type and brand of PV modules. The PV panel power generation model (object) in EnergyPlus has also been used in a lot of research studies to simulate the operation and enhance the efficiency. Equivalent one-diode and Sandia models EnergyPlus were used to model PV arrays on different sites and the results are validated with experimental data in [41]. Saber et al. [42] applied all these three types of models to simulate PV systems. They results from these three models are compared and validated against the measured data. Simple PV model was used in [43] to simulate PV panel system on vertical and horizontal building surfaces. The Equivalent One-Diode Model was employed to simulate a building integrated PV panel systems for a multistory residential buildings in [44]. The Sandia model was used by Stefanvic et al. [45] to simulate and choose the operation strategy of PV panel to achieve a Net-Zero building.

There is no ice tank thermal storage model in TRNSYS. However, it provides a lot of PV panel and battery models (Types), and they have been used in many studies for PV panel power generation estimation under different conditions, such as different location, different array inclination, different orientation [46-52]. Different battery models (Types) are also provided in

TRNSYS, such as shepherd simple lead-acid model, shepherd model with hyman equation, and shepherd model with gassing current effects model. These battery models are mostly used in conjunction with PV panel model. In [46] and [51], the default simple shepherd lead-acid battery Type was used and connected to the PV panel model.

No studies about modeling ice tank thermal storage using Modelica have been found in the existing literature. Just one paper has been found in using Modelica to simulate the PV panel power generation system [53]. In this paper, detailed representations of the PV system and the building thermal response models, and air–water heat pump are implemented in Modelica.

Even though these elaborate simulation tools are effective and accurate, they require detailed information and parameters of buildings, energy systems and outside weather conditions. These parameters, however, are always difficult to obtain, and even sometimes are not available. What's even more challenging, creating these white box models normally requires expert work, and the calculation is extremely time-consuming, which is the major barrier for white box building models to be used in on-line model based control and operation.

1.2.1.4 White Box Model for Building Control and Operation

White box models, including whole building and component models have been applied in building control and operation studies. Although taking more time for developing and calculation than black box and gray box models, white box models are typically have better accuracy over a wide range of operating condition. Hence due to the improvement of computer's calculation ability, they have been used for building control and operation in this decade. It is also expected that more and more white box models (especially whole building ones) are developed during the modern building design process, which can later be used for building control and operation purposes. The recent studies using white box for building operation purpose are reviewed in this section.

A Matlab-EnergyPlus MPC environment testbed has been created in [54]. The procedure of this MPC is illustrated in Figure 1-5. Before the simulation engine read in all these files, candidate system parameters which need to be optimized would be written into the idf file, replacing the original data. EnergyPlus output results would then be evaluated within the MATLAB optimization module. Based upon the objective cost function, the operation parameters would be updated and written in to "idf" file again. This procedure will be repeated until the operation criterions are satisfied. Using this environment, the authors created a generalized linear model to extract near-optimal heuristics which can be implemented in to the building control systems. The logic "extraction" is developing a generalized linear model in the form of multi-logistic regression to build simplified decision models that can mimic the optimizer results.

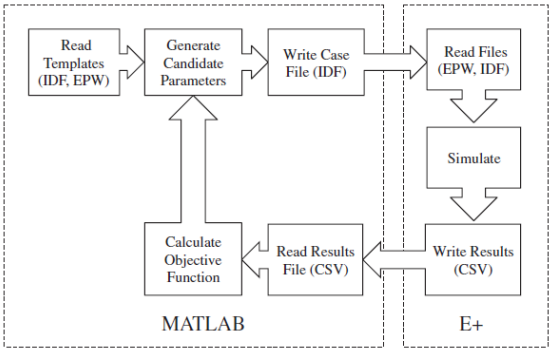


Figure 1-5. Matlab-EnergyPlus MPC procedure [54]

Corbin et al. [55] further utilized this Matlab-EnergyPlus MPC environment and incorporated it with a particle swarm optimizer to predict optimal building control strategies. A meta-heuristic search technique, particle swarm optimization (PSO), has been adopted to search for near-optimal candidate parameters. In this paper, an on-line building operation optimization scheme has also been developed and demonstrated through a server-client framework. This framework allows optimized control strategies to be exchanged between the optimization model and the building automation system (BAS). This on-line optimization environment was also applied in a DOE benchmark building EnergyPlus models. The results showed 5% cost saving just by optimizing

the hourly cooling setpoints in a large office building model, and 54% energy saving by optimizing hourly supply water temperature in a small office building model. Yoon et al. [56] developed a demand response strategy for a residential building based on the dynamic electricity price. Performance of demand response controller (DRC) for homes is modeled by EnergyPlus. The simulation results show that DRC reduces significantly HVAC loads during peak time.

An important control optimization software environment, Genopt, was developed by Wetter [57], which can iteratively execute any simulation program based on plain text input/output files, such as EnergyPlus until an optimal solution is found. Genopt was used by Coffey et al.[58] to incorporate a modified genetic algorithm model predictive control with an EnergyPlus model to study the temperature control optimization in office buildings and its effect building energy demand. Rackes and Waring [59] used multi-objective optimization to determine the building dynamic ventilation strategies to save energy and improve indoor air quality, using EnergyPlus and Genopt. Seo et al. [60] also used EnergyPlus and Genopt to develop an optima lighting and daylighting control strategy. The new lighting and daylighting control strategy can be incorporated in an energy management and control system (EMCS) in EnergyPlus to operate and control lighting fixtures in any indoor space. EnergyPlus is one of the most comprehensive white box models is EnergyPlus. Unfortunately, thousands of parameters in EnergyPlus models need to be identified, so it is very time consuming to create an EnergyPlus model. What's worse, the simulation speed is relative low and not suitable to be used in on-line MPC. Therefore, Cole et al. [61] used OpenStudio to reduce the EnergyPlus model and succeed in applying this reduced model in a MPC model.

A simulation-based multi-objective optimization framework based on a combination of TRNSYS, Genopt, and MATLAB has been developed to optimize the energy saving, thermal comfort and cost of a residential building [16]. This framework is mainly for building retrofits,

which can simulate various retrofit actions and choose the optimal combination. Al-ajmi and Hanby [17] simulated the energy consumption for a commercial building in Kuwait using TRNSYS. Several parametric studies were conducted to enable the sensitivity analysis for the building energy efficient, such as analysis for building envelope insulation, infiltration, and window type. In order to determine the heating system, Goncalves et al. [18] implemented a building model in TRNSYS with 8 different heating options. The overall energy and exergy performance of these 8 space heating options is compared for different outdoor environmental conditions. Khandelwal et al. [19] explored the energy saving potential of regenerative evaporative cooling in a three-floor library building using TRNSYS. The regenerative evaporative cooler is coupled to the existing water chiller to examine the energy saving and comfort level in comparison to the existing fan coil system. Carpenter et al. [23] investigated the energy saving potential of passive thermal mass management by circulating water through a piping system located in the building walls or ceiling. In the TRNSYS simulation of a reference building in Atlanta, GA, around 7% cooling load reduction has been achieved from the circulating water in the walls, and 11% reduction from the circulating water in ceiling.

Tomažič et al. [29] developed a Matlab/Simulink and Modelica simulator for indoor environment dynamic control, where the thermal model of the simulator was developed in Modelica, and illuminance model was developed and parameterized in MATLAB. In [30] and [31], two different module for building thermal behaviors have been developed in Modelica. Comparing to the building models in TRNSYS, these two models used object-oriented approach.

1.2.2 Black Box Simulation Models

Black box model is also known as data driven model. Typically, statistical models are applied to capture the correlation between building energy consumption and key operation data. This type

of models needs on-site measurements over a certain period of time to train the models to be able to predict the building operation under different conditions. These black box models are also widely applied in existing studies to determine building control strategies to reduce energy consumption and energy cost. In this section, different building energy black box models using different methods, such as regression, artificial neural network (ANN), and support vector machine for regression (SVR) are introduced.

1.2.2.1 Black Box Simulation Methods

Multivariable regression, ANN, and SVR are the most common black box methods used in building energy forecasting. The principle of multivariable linear regression is to predict output as a linear combination of the multiple input variables. Artificial neural networks are generally presented as systems of interconnected "neurons" which can compute values from inputs. SVM for regression is generated by support vector classification to find an optimal generalization of the training data set.

There are different extended forms of Multivariable regression, such as autoregressive moving average (ARMA) models, and autoregressive exogenous (ARX). They have been applied to forecast the building load [62-65] and energy consumption [66-70] starting from 1980s. ARX model is reported as the most efficient regression model to simulate building energy and has been adopted in a lot of studies [71-76].

Artificial neural network (ANN) is another popular method in building energy modeling for building operation purpose. In particular, the adaptability of ANN models through a self-tuning process, which is different from mathematical models such as regression models, makes accurate decisions under the disturbance. Kalogirou published many papers, including a review paper, about using ANN for building energy estimation [77, 78].

SVM is increasingly used in research and industry due to its highly effective model in solving non-linear problems [79]. So far it has been widely used in various analyses such as regression, classification and non-linear function approximation.

1.2.2.2 Black Box Model for Whole Building Simulation

ARX model was developed and implemented to predict the one hour ahead building load in [72]. This predictive model is applied on several different DOE benchmark buildings [15] to choose the building control strategies. Aydinalp, et al. studied the annual electricity consumption in residential sector of appliances, lighting and cooling in [80] and of space heating (SH) and domestic hot water in [81].

Yang et al. [82] developed and validated adaptive ANN model for on-line building energy prediction. Kwok and Lee studied the influence of the occupancy on the cooling load using ANN [83]. They compared three different neural networks to predict the total building cooling load: a network for weather conditions, a second one for occupancy area, and a third one occupant behavior. Moon and Kim proposed a thermal comfort control method using an ANN based model to enhance thermal comfort in residential buildings [84]. Yokoyama et al. [85] used a back propagation neural network to predict cooling demand in a building. The energy demand was predicted using its measured or predicted values as well as the predicted values of air temperature and relative humidity. This approach was applied to the prediction of the cooling demand in a building used for a bench mark test of a variety of prediction methods to validate its effectiveness. Even though ANN has been widely successfully applied in different studies, it is still limited by its lack of interpretability and high quality training data requirement.

Dong et al. [86] were the first to use SVR for building energy consumption prediction. The input variables are the mean outdoor dry-bulb temperature, the relative humidity and the global solar radiation. Li et al. [87] used the SVR with radial basis function as kernel function for the

prediction of hourly cooling demand in Guangzhou, China. Zhao and Magoules [88] used SVR with Gaussian kernel predicted the energy consumption of multiple buildings. Same model was also applied by Lai et al. [89] to forecast the electrical consumption in residential sector of Tohoku, Japan. Hou and Lian [90] compared ARIMA model and SVM model in cooling load predicting, and the results showed that SVM performance better than ARIMA.

1.2.2.3 Black Box Model for Building Component Simulation

Bacher et al. [91] applied regular auto-regression and ARX to on-line forecast the solar power production from PV systems. Giraud et al. [92] analyzed performance of PV and battery system under the effects of a passing cloud using ANN. An adaptive ANN is applied to model a stand-alone PV system under variable climatic conditions. Levenberg–Marquardt algorithm and infinite impulse response (IIR) filter have been used to accelerate the convergence of the network [93]. Kalogirou et al. [94] used ANN models to predict the performance of large solar systems. They concluded that the ANN effectively predicts the daily energy performance of the system; the statistical R^2-value obtained for the training and validation data sets was better than 0.95 and 0.96. ANN has also been used to model the performance of thermal energy storage system [95] and electrical energy storage system [96]. SVM has also been used in building component modeling. Zeng and Qiao predicted the short term solar power generation based on SVM in [97], where the prediction results from SVM based model were better than those from auto-regression model and ANN model. Bouzerdoum et al. [98] developed a hybrid model for small scale grid connected PV plant power forecasting model based on ARIMA and SVM. A battery performance estimation model was developed by Anton et al. [99] using SVM.

1.2.2.4 Black Box Model for Building Control and Operation

In recent years, more and more studies apply black box models for building control and operation purposes. Table 1-1 summarizes the studies of black box model for building control.

Black box models are easy to build and computationally efficient, however, such models often require long training period and are bounded to building operating conditions that they are trained for which sometimes can cause huge forecasting error when training data does not cover all the forecasting range.

Table 1-1. Black box model for building control and operation

	HVAC system	**Energy generation**	**Energy storage**
ARIMA, ARX	[71, 76, 100, 101]	[91, 102]	[39, 103]
ANN	[84, 104]	[92, 93, 102, 105]	[92, 95, 96]
SVM	[86, 106, 107]	[102, 108]	[109]

1.2.3 Grey Box Simulation Models

Grey box models are hybrid models that use simplified physical descriptions to simulate the behavior of building energy systems. Using the simplified physical models reduces the requirement of training data sets, and calculation time. In a gray box model development process, model coefficients are identified based on the operation data using statistics or parameter identification methods. A unique three-step process was developed for developing and training the gray models in a previous study [110]:

1. Develop the simplified physics models for building;

2. Bounds on physical parameters are estimated from a rough description of the building geometry and materials;
3. A rough parameter identification algorithm is used to determine estimates of model parameters.

Typical simplified physics based models include thermoelectricity analogy structure for building envelope heat balance simulation [110], and lumped parameter models for energy devices [111]. The bound of parameters are directly determined by the property of building systems, such as the heat transfer coefficient and thermal capacity of building envelope, energy device performance coefficient, and so on. Last, there are numerous methods for parameter determination, such as, regression methods, maximum likelihood method, and optimization prediction error approach, and etc.

In this section, the application of grey box models for whole building, building component simulation and building operation is discussed.

1.2.3.1 Grey Box Models for Building Energy Estimation

In building envelope heat balance modeling, resistance and capacitance network (RC) model is the most commonly used method, where capacitors (C) represented the thermal capacitance, and resistors (R) between the nodes represented the thermal resistances of building envelope and indoor mass. The benefit of this RC model is its physics representation and computation efficiency.

A RC model structure (Figure 1-6) was developed to model and predict building cooling load in [110]. The values of resistances and capacitances were determined by “nonlinear regression” method of on-site measured operation data.

Wang and Xu developed a similar simplified whole building level energy model in [112]. However, different model structure has been applied in modeling external and internal envelopes due to their different dynamics characteristics. The parameters in this network were determined by using genetic algorithm (GA) based searching methods. Zhou et al. [113] developed an on-line next day building load prediction model for building energy efficient control, by using the grey box building energy model developed in [112] and a weather forecasting model. The results of this study show that this building load prediction method is suitable for the on-line prediction of building load for the coming day and coming hours, whose mean of absolute percentage error is below 8%. Besides these studies, there are a lot of studies employing the RC network to estimate the building energy consumption [113-119].

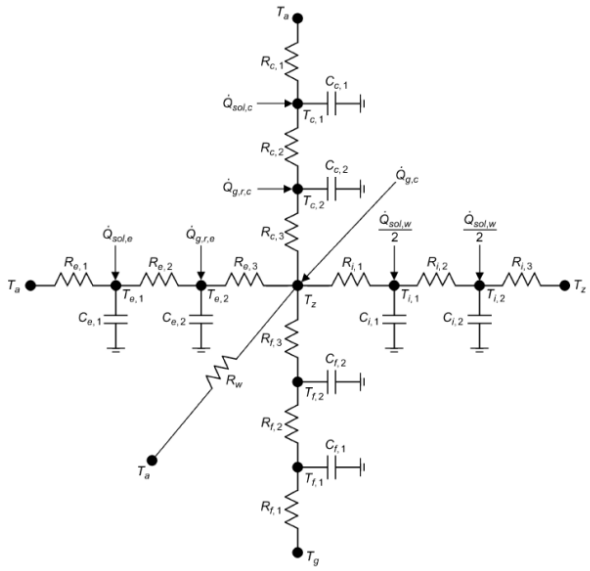


Figure 1-6. Thermal network for overall building model [110]

1.2.3.2 Grey Box Models for Building Component Simulation

Grey box models are widely used in building component, such as HVAC, PV, and battery systems, modeling, because the Physics equations of these systems are relative clear but the parameters are varying. Hence, general simplified Physics models with different parameters are

able to model the behaviors of different types of these systems. Therefore, there are many studies applying grey box models to simulate different building components, for example, building HVAC systems [120-124], ice tank thermal energy storage systems [38, 125-129], PV and battery systems [130-134]. Most of these grey box models have been used in model based control and operation for building energy efficiency improving. The details of using grey box modeling approach to simulate the performance of building energy components will be discussed in Section 5.1, when those models are developed in the dissertation.

1.2.3.3 Grey Box Models for Building Control and Operation

Comparing to white box and black box models, grey box model is more commonly used in model based control and operation studies, due to its better performance in modeling speed then white box model and accuracy than black box model.

Grey box model is widely used in building temperature control to fully utilize the building passive thermal mass storage to save energy and maintain indoor thermal comfort, such as [115, 117, 125, 135-140]. The general approach for these studies is using simplified physics models to estimate the building heating/cooling load at certain temperature setpoints, and HVAC device models, such chiller model, will be used to calculate the building energy consumption using the building load. Finally, an optimization model will be developed to determine the optimal temperature setpoints:

$$J(t_i) = \min_{t_i} \sum_{i=1}^{n} C_{i,e} P_{i,e} + C_{i,d} P_{i,d} \qquad \text{Eq. 1.1}$$

where, $C_{i,e}$ is the energy consumption, $P_{i,e}$ is the price of energy, $C_{i,d}$ is the energy demand, and $P_{i,d}$ is the demand cost.

PV panel energy generation and ice tank energy storage provide more opportunities to dispatch building energy peak demand and reduce energy cost than thermal mass energy storage.

In the studies of model predictive control for these building devices, grey box models has also been widely used, such as for PV panel power generation [132, 133, 141-143] and for ice tank thermal storage [39, 125, 126, 144-148].

Traditionally, the control scheme for PV systems is “PV Priority” control scheme. In this control scheme, electricity generated from PV panel will power the load and then the excess power, if any, will charge the battery. The PV priority control is very simple and easy to implement, but it is not optimal. Hence, a lot of complex control strategies have been proposed and tested in this decade. The basic logic is to express the desired system behavior in an objective cost function, and then optimize this cost function over a time horizon using some optimization method. In [132], an adaptive optimal control scheme was developed and tested for a grid-independent photovoltaic system using Q-learning algorithm, with a cost function placing more weight on meeting a critical base load than on those non-critical loads exceeding the base load. Nevertheless, Q-learning algorithm, as a model-free learning method, requires much on-site data in order to form solid estimates of its parameters.

Similar to the control of PV panel system, model based control for ice tank is also utilize the ice tank model to determine the optimal control strategy. In [147], two thermal energy storage models were developed and combined with a dynamic building model and a utility rate module to investigate the energy saving potential of ice tank thermal storage. A cost optimization approach to minimize total utility bill either with or without demand charges based on a time-of-use electricity rate, using direct search, gradient-based methods, and dynamic programming method, was incorporated into the simulation model. 17% to 27% cost saving was achieved in a small commercial building in Iowa, comparing to the traditional “storage-priority” control schemes. In [148], Vetterli and Benz investigated an optimal ice storage cooling system under various electricity tariff schemes Mixed integer linear programming method was implemented to linearize

and solve the linear optimization problem of the building energy model, ice storage system model and dedicated vapor compression chiller model. Approximately 8% cost reduction was achieved from this optimal design and control of ice tank storage system in a cast study in this paper.

Comparing to white box and black box models, grey model has its own benefits in application in model based control and optimization, due to its advantages in simulation accuracy and speed. However, even creating a simplified physics based model is often challenging and time consuming. The model structure and order are usually determined based on the prior knowledge or trial and error process, which cannot guarantee the accuracy and usually takes too much time. Therefore, looking at the problem in system identification direction provides more promising opportunity to develop the model based on the characters of the system.

1.2.4 System Identification Methodologies for Building Energy Forecasting

As discussed earlier, compared with white box model, black box and grey box models have the advantage of simplicity and computational efficiency. However, generating black box and grey box models from data that are collected either during the building design process or during a regular building operation process is a time consuming process. The model accuracy is also strongly affected by the diversity of the operating ranges/conditions that the collected data cover. In another word, the model parameters/structures are determined in a passive manner. On the contrary, system identification is a process of developing or improving a mathematical representation of a physical system using data that are collected from a designed operation or an experiment, in an active manner [149]. Although system identification techniques have been used in other engineering applications [150-154], there are only limited applications of system identification techniques in building modeling.

In a system identification process, four steps are usually involved: 1) system prior information collection through active experiment; 2) system identification structure selection; 3) system order determination; 4) system parameter identification [149]. The detailed procedure of each step will be introduced in section 4. In this section the building energy modeling studies using system identification approach will be reviewed.

As introduced, only a handful studies have applied system identification in building energy modeling. Privara et al. [155] proposed an approach combining the EnergyPlus model and a subspace system identification model to forecast building performance. Since building regular operation data range is relative narrow and it lacks of high frequency information, a MATLAB-BCVTB-EnergyPlus testbed was developed for building excitation, system identification and building performance forecasting. Pseudo-random binary signals, sum of sinusoid signals and multilevel pseudo-random signals were used to excite the building system by updating temperature setpoints to get good quality building operation data for model training. Then a subspace model in MATLAB N4SID toolbox was utilized for system identification and operation forecasting.

This same building system exciting method was also applied in another publication [156]. Similar to all the system identification studies, besides the model structure selection, the system order and the Hankel matrix size are important factors for model accuracy and calculation speed. In this study, parametric testing studies were conducted to determine these two factors. From their testing and validating results, 18^{th} order subspace model turned out to be the best choice, considering both its simplicity and sufficient precision. Considering the system order and richness input signal data, finally 40 was selected for the size of Hankel matrices for one subsystem. Although parametric testing methods were useful to determine the system order and Hankel matrix rank, systematic approach for model structure selection, order determination still have not

provided. Most of the existing studies just tried the trial-and error method to determined model structure and order.

Multi-step ahead identification with least square method were developed for building energy modeling in [157]. Singular value decomposition (SVD) decomposition was used to determine the order of the multiple-Input Multiple-Output building heating energy prediction system. A case study was presented using this prediction system. 600 data points representing 7 days from Jan. 3rd to Jan 10th 2011 were used to train the system. Based on the training data and SVD results, the system was chose to be a 5th order model. A MPC model was then developed based on this prediction system in control of building temperature setpoints.

As an active system modeling process, system excitation signal is crucial to system identification models' accuracy and robustness [149]. Different system excitation strategies, such as Pseudo-Random Binary signal, Pseudo-Radom Sequences, Multi-sine signal, have been discussed and applied in [150, 158, 159] on-linear process systems in Mechanical and Chemical process. Excitation signals' constraints and guidelines to ensure the signals contain enough frequency information to meet the identification requirements have been established and tested. For example, a Multi-sine signal, as shown in Eq. 1.2, was designed and applied in [159].

$$u_s(k) = \lambda \sum_{i=1}^{n_s} \sqrt{2\alpha_i}\, cos\,(\omega_i kT + \phi_i) \qquad \text{Eq. 1.2}$$

$$\frac{1}{\beta_s \tau_{dom}^H} \leq \omega \leq \frac{\alpha_s}{\tau_{dom}^L} \qquad \text{Eq. 1.3}$$

Where λ is the scaling factor, T is the sampling time, n_s is the number of harmonics to reside within the frequency limits, and ϕ_i is the phases lag. The frequency ω is constrained by Eq. 1.2. Where τ_{dom}^H and τ_{dom}^L correspond to the high and low estimates of the dominant time constant of

the system. α_s and β_s are user-decisions on high and low frequency content based on identification requirement. This excitation signal generation method has been used in [155].

Another Pseudo-random Binary Sequence (PRBS) excitation signal (Eq. 1.4) for building temperature setpoints was generated and applied in a building energy modeling study discussed previously[76].

$$T_{sp,i}(k)\begin{cases} 21, & excited\ zone, \text{PRBS} = 0 \\ 25, & excited\ zone, \text{PRBS} = 1 \\ 25, & nonexcited\ zones \end{cases} \quad \text{Eq. 1.4}$$

System identifiability is another very important factor to system identification accuracy and efficiency. It is affected by the input data, excitation signals and system model structure. A effective theoretical study about system structure and local identifiability based on excitation signal inputs and system measurements has been published in [160]. This study presented a building data dependent identification algorithm to calculate the numerical identifiability for high order RC model, by checking the rank of its Hankel matrix. This algorithm is a closed-loop "active identification" structure which can be used to improve the experimental design for better training data quality. Multi-sinusoidal and random Gaussian excitation inputs were injected into the model as excitation signals to identify those resistances and capacitances based on building operation measurements and excitation signals.

Although studies started to apply system identification approach to model and forecast building energy performance, there is a lack of systematic analysis about the system structure selection, system order determination and the relationship between excitation data quality and system identifiability and efficiency. How to generation appropriate excitation signals, training data, and determine model structure, order with considering the characteristics of building energy systems is an urgent research topic for system identification in building energy modeling

1.2.5 Data Fusion Techniques for Building Energy Forecasting

In previous sections, different building energy modeling approaches have been reviewed. They all have been successfully applied in several studies with acceptable performance. However, most of the models, even with system identification process, may not be accurate after a while when the system changes over time. Therefore data fusion techniques can play an important role here to update and calibrate the model based on real measurements to improve its performance.

Data fusion is a process of integrating multiple data for a real world system into a more accurate and robust representation. Castanedo [161] published a review paper for data fusion techniques, where the popular data fusion methods and techniques, such as Kalman filter, extended Kalman filter were reviewed in this paper. The details about these techniques are introduced in section 6.1.

In building energy forecasting area, data fusion techniques have also started to be used to improve the model accuracy and robustness. An integrated 3R2C and EKF (Extended Kalman Filter) model was developed to estimate the building energy consumption in [162]. This RC and EKF approach has also been utilized in [163], where self-adaptive thermal building model was developed based on a 1R1C model and an EKF. This 1R1C and EKF model can reduce parameter identification time and calculation burden. 1R1C sacrificed the estimation accuracy which would be made up by incorporating real measurements through EKF. Thus the balance between process noise covariance and measurement noise covariance determination is crucial to the model accuracy and robustness. Unfortunately, neither of these two studies has discussed the details about the determination of process noise covariance and measurement noise covariance, and even the correlation between these two noises.

1.2.6 Co-simulation Tools: Connecting Different Models

Building energy modeling for control and operation purpose often involved different models in different environment. On the other hand, a simulation testbed, including different models in different environment, is also need to validate the on-line models and the operation strategies. Hence, co-simulation tools are needed to combine different simulation models in different environment. The general information of two most widely used co-simulation tools and their application in building energy modeling will be introduced.

1.2.6.1 MLE+

MLE+ [164] is a Matlab toolbox for co-simulation with the whole-building energy simulator EnergyPlus. It is free under an open source license from University of Pennsylvania. MLE+ is particularly designed for controller design, where the energy simulation is carried out by EnergyPlus while the controller is designed and implemented in Matlab or Simulink, and simulation-based optimization, where a non-linear optimizer, e.g. one in the Matlab Global Optimization Toolbox, can be used to find optimal parameters or control sequences of the building system, by considering the building as a black-box whose execution is performed by simulations by EnergyPlus.

1.2.6.2 Building Controls Virtual Test Bed (BCVTB)

BCVTB [165] is another freely available, open source co-simulation software, provided by the Lawrence Berkeley National Lab, and is based on the Ptolemy II software environment that has been developed by the University of California at Berkeley. BCVTB is able to couple several different simulation programs, including EnergyPlus, Matlab/Simulink, Modelica, TRNSYS, and BACnet which allows data exchange between simulation programs and real Building Automation System (BAS). For example, given EnergyPlus-simulated environmental parameters at a certain

time step, the BCVTB could provide this information as input to a behavioral algorithm implemented in Matlab or TRNSYS, which would compute a corresponding building energy devices' operation and pass their operation data back into EnergyPlus or Matlab to adjust relevant parameters or control signals for the next time step.

Based on BCVTB, a framework for simulation based real time building performance assessment has been developed in [166]. In this framework building energy simulation models in EnergyPlus, operation models in MATLAB, and real building energy management and control system in BACnet were connected together through BCVTB. Ma et al. [76] proposed and demonstrated an economic model predictive control (MPC) technique to reduce energy and demand. The economic model was implemented in MATLAB to generate building operation signals for building EnergyPlus model. About 25.3% energy saving and 28.5% cost saving were achieved by this MPC in a single story commercial building located in Chicago, Illinois. An EnergyPlus-Simulink co-simulation framework has been developed to test a new HVAC control method based on BCVTB [167]. The new HVAC control scheme uses PMV of each occupant as feedback and provides the opportunity to act on their own comfort level by signaling a thermal sensation in a personal user interface.

1.3 Research Need and Scope

1.3.1 Research Needs

As discussed in Section 1.1, a building cluster simulation testbed is needed to simulate the behaviors of multiple buildings, as well as building energy devices in order to develop and assess the control and operation strategies. However, all existing building simulation testbeds focus on individual building and its related devices. Therefore, the first objective of this project is to develop a building cluster emulator that is capable of simulating realistic energy behaviors of a

cluster of buildings and their energy generation and storage devices, such as PV panels, battery devices, thermal storage tank etc.

This emulator is for developing and accessing building cluster operation strategies by providing real-world-like building cluster operation data. However, adopting such detailed emulators in an online building cluster control and operation strategies will be too expensive and computationally inefficient. As indicated in the literature, extensive studies exist in building energy forecasting area. Yet each of white box, gray box, and black box modeling approaches has its own shortcoming, which prevents it from delivering a satisfactory performance in terms of accuracy, computational efficiency, and model development cost for online building cluster energy forecasting. Meanwhile, system identification and data fusion methods have shown great promises to be combined with traditional building modeling approaches to deliver models that have high fidelity, yet are also computationally and costly efficient.

Thus in order to determine the optimal operation strategies, this dissertation proposes to apply system identification and data fusion techniques to develop high fidelity on-line building cluster energy forecasting models. Therefore, the two main researching objectives in this project are to a) develop a building cluster emulator for building cluster control and operation strategy assessment and b) develop high fidelity on-line building cluster energy forecasting models to be used in determining building cluster control and operation optimization strategies.

1.3.2 Research Scope

In support of the two research objective, three primary tasks are proposed here.

Task 1 Building Cluster Emulator Testbed Development (CHAPTER 2)

Using existing building and building component simulation models and software, a building cluster emulator needs to be developed to provide real-world-like operation data, and to be able to be connected with real building EMCS system

Task 2 Building Cluster Energy Forecasting Models Development and Validation (CHAPTER 3, 4, 5 and 6)

Task 2.1 System Identification Feasibility Study (CHAPTER 3)

In this task, the feasibility of using system identification techniques to develop a while building energy forecasting model is studied. Building energy operation data from cluster emulator testbed are used to develop and validated the system identification model. A strategy to excite the building energy systems for model training is also demonstrated.

Task 2.2 Develop General Building Energy System Identification Methodology (CHAPTER 4)

Based on the findings from the feasibility study, an overall system identification methodology for building energy forecasting is developed here.

Task 2.3 Develop Data Fusion Methdology (CHAPTER 6)

The feasibility of using Kalman filter extended Kalman filter to further improve the online building energy forecasting model is studied here. Based on the feasibility study, a framework of applying data fusion for building energy forecasting is developed.

Task 2.4 On-line Building Cluster Energy Forecasting: Other Modeling Approaches (CHAPTER 5)

Although a new System identification modeling methodology is proposed in Chapter 3 and 4, existing gray box models for alternative energy components, energy storage components and the whole building (RC models) are examined here. Models for other building energy devices simulation are also development in this chapter

Task 3 Whole Building System Identification Experiment: Model Refinement and Validation (CHAPTER 7)

A series of real building experiments that are designed to evaluate, validate, and improve the proposed online building energy forecasting models

2. CHAPTER 2 Building Cluster Emulator Testbed Development

2.1 Background and Motivation

The emerging and improvement of smart grids, net–zero energy buildings, and advanced building energy demand response technologies continuously drive the needs for better building control and operation strategies and building energy simulation testbeds. However, current building operation strategies and simulation testbeds are all developed assuming that buildings do not share energy resources. In this study, it is envisioned that similar to micro-communities in a human society, neighboring buildings will have the tendency to form a cluster, an open cyber-physical system to exploit the economic opportunities provided by a smart grid, distributed power generation, and storage devices. Through optimized demand management, these building clusters will then reduce overall primary energy consumption and peak time electricity demand, and be more resilient to power disruptions. This building cluster concept will be able to fundamentally transform the energy industry by shifting expensive on-site energy generation aimed at creating NetZero buildings one-at-a-time to an autonomous and adaptive system of buildings aimed at NetZero clusters.

To better develop the operation strategies for such net-zero energy building clusters, this task aims at developing a building cluster emulator that is capable of simulating realistic energy behaviors of a cluster of buildings and their energy generation and storage devices, such as PV panels, battery devices, thermal storage tank etc. The emulator is able to interact with simulated control and operation strategies, as well as to connect to real-world building control system through BACnet interface in the future. Multiple simulation environments are utilized in developing the emulator: each building in this cluster as well as a shared thermal storage device is modeled in the EnergyPlus environment. TRNSYS is utilized to simulate electrical power

generation and storage systems. All models are interconnected through a Building Controls Virtual Test Bed (BCVTB) environment, which is also used to connect the emulator with a commonly used computation environment, as well as with a BACnet interface in the future. This task will provide the foundation for the following research tasks:

1. This emulator will simulate the building cluster operation and provide real-world-like operation measurements;
2. It can be used for numerical experiments to provide training and validation data for on-line building energy forecasting model development;
3. It will be used to assess the efficacy of any proposed operation strategies. A proof-of concept demonstration is also conducted in this study.

2.2 Building Cluster Emulator Overall Design and Operation

2.2.1 Emulator Overall Connection Design

The overall emulator design and connection are illustrated in Figure 2-1. There are 4 modules in the emulator: Building module, Ice tank module, PV-Battery module, and Control module. The buildings in the Building module share the PV-Battery and the Ice tank module. Building, energy generation, energy storage and operation models are developed in different software. EnergyPlus is chosen to simulate the buildings and ice tank thermal storage devices, because it is widely used and validated to provide detailed simulation results at a minimum one minute time step. PV panel power generation and battery system is modeled in TRNSYS by its default Types. Control module, which can utilize any control/operation strategies that need to testbed resides in MATLAB [168]. BCVTB is served as a middleware to connect EnergyPlus, TRNSYS and MATLAB.

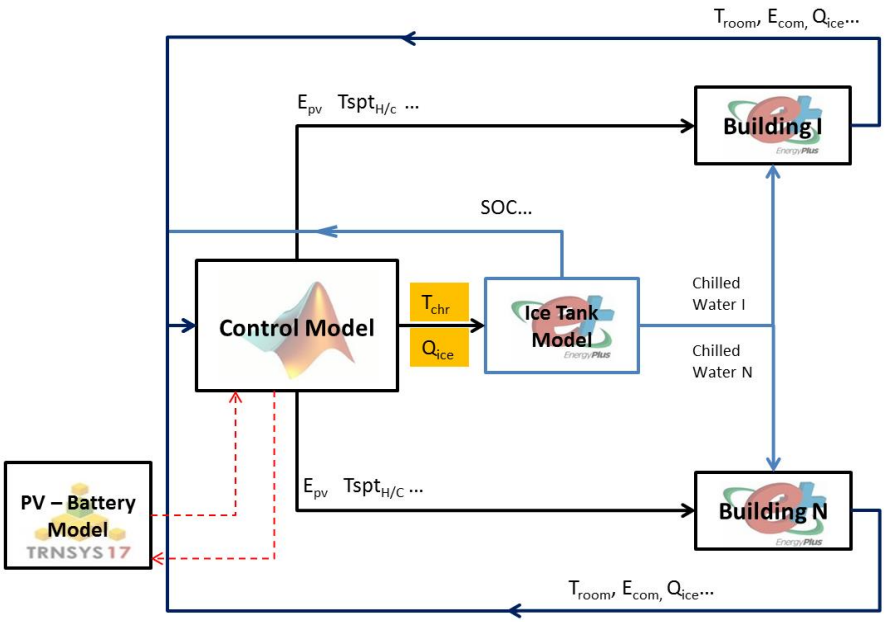


Figure 2-1. Emulator operation control diagram

2.2.2 Emulator Operation Design

The emulator can apply control and operation strategies provided from MATLAB (and BACnet in the future) and simulate the operation of various systems and devices included in the cluster. The inputs and the major outputs of the emulator are summarized in Table 2-1. At each simulation time step, the emulator can provide real-world-like "noisy" and "noise-free" measurements based on user requirements, such as energy generation and energy storage devices' performance data, etc. The control signals sent from the control module are building heating and cooling setpoints (Tseph, Tsepc), operation energy generation and storage (Si, Sb, Spv), electricity purchasing and selling (Ep, Es), out of which the first two are sent to the Building module, Si is sent to the Ice tank module, and Sb and Spv are sent to PV-Battery module, Eb and Es are sent to power grid module. Meanwhile, building measurements (Eb, Tzone, etc.) are output from Building module, ice tank measurements (SOC_i, Tchlw, Mchlw, and Eice) are output from Ice tank module, PV-Battery measurements (Epv, Ebat, and SOC_b) are output from PV-Battery

module, Ptou, Pc and Pe are output from power grid module. All these outputs will be transferred to control model to assess the control strategies at each time step and a set of new control signals will then be sent to the cluster.

Table 2-1. Building cluster input-output summary

Sub-Module	Input	Major Output
Building model	Tseph (Heating setpoint) Tsepc (Cooling setpoint)	Eb (Building energy consumption), Edc (Dedicated chiller energy consumption), Ebc (Base chiller energy consumption), Tzone (Building zone temperature), Hzone (Building zone humidity)
Ice tank	Si (State of ice tank storage)	SOC_i (ice tank state of charge), Eice (charging and discharging rate), Tchlw (chilled water in/out temperature), Mchlw (chilled water in/out temperature flow rate)
Battery model	Sb (State of battery)	SOC_b (battery state of battery), Ebat (charging and discharging rate), Ibat (charging and discharging current)
PV model	Spv (State of PV panel)	Epv (Power generation)
Power grid model	Ep (Power buying from power grid) Es (Power selling to power grid)	Ptou(Time-of-use price), Pc(electricity costs), Pe(electricity earnings)

2.3 Building Cluster Emulator Development

2.3.1 Building Module Development

Although the testbed can include any number and any type of EnergyPlus building model, in this paper, EnergyPlus models for two different buildings are identified. Two typical medium-size office building models in Philadelphia which were developed in a previous study are selected for this cluster emulator [169]. This first building is a one-story, 5,000 square feet (464.5 square meter) commercial building and the other one is a three-story, 15,000 square feet (1393.55 square meter) commercial building. The window to wall ratio for both buildings is approximately 0.29. The windows are of various single and double pane construction with 0.118 inch (3 mm) and 0.236 inch (6 mm) glass and either 0.236 inch (6 mm) or 0.512 inch (13 mm) argon or air gap. The U-factors of the windows are 1.0 Btu/hr-ft^2-°F (0.173 W/m^2-K) and 0.5 Btu/hr-ft^2-°F (0.086 W/m^2-K), respectively. Both buildings have deck roofs with R-15 insulation (solar absorptivity of 0.9). The summary of the mechanical systems are tabulated in Table 2-2. The first building

system is a single duct constant-air-volume (CAV) roof-top units (RTUs) system and the other one is a single duct CAV air-handling units (AHUs) system.

The summary of their mechanical systems are tabulated in Table 2-2. The first building system is single duct constant-air-volume (CAV) roof-top units (RTUs) and the other on is single duct CAV air-handling units (AHUs).

Table 2-2. Building mechanical systems

	Building I	**Building II**
System	3 CAV, RTUs	3 CAV, AHUs
Main Cooling Coil	DX, COP 3	Chilled water
Main Heating Coil	Hot water	Hot water
Zone Reheat	Hot water	Eclectic
Heat Plant	Central Boiler	Central Boiler

2.3.2 PV-Battery Module Development

PV panel power generation and battery electricity storage are also modeled in this emulator. The development of the shared PV panel-Battery system module will be introduced in this section. As shown in Figure 2-2, the PV-Battery system module contains a PV panel model, a battery model, a power grid model, a local operation controller model and two inverter models.

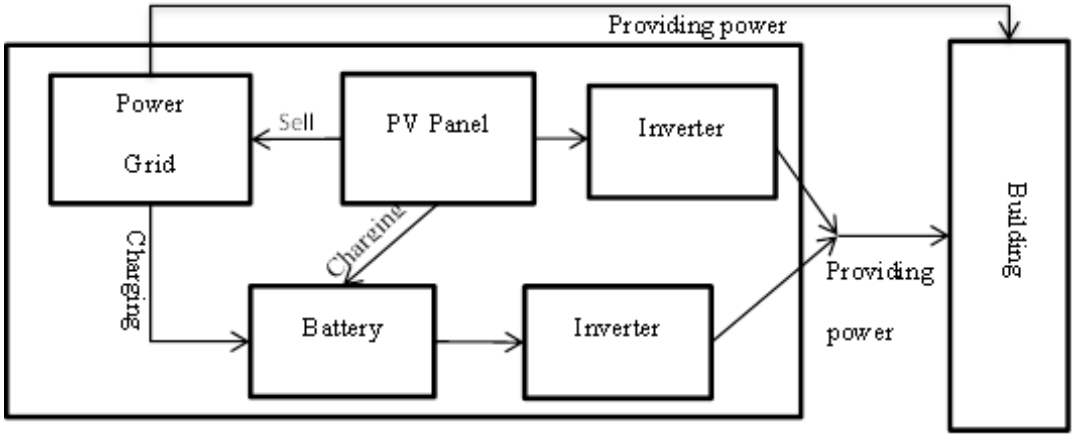


Figure 2-2. PV panel-battery system configuration

All these components are modeled individually and interconnected in TRNSYS. The PV panel is modeled by Type 194 in TRNSYS, which is based on the diode equivalent circuit model to calculate the power generation rate, outlet current, and voltage etc.[170]:

$$I = I_L - I_0 \left[exp\left(\frac{q}{\gamma k T_c} (V + I R_s) \right) - 1 \right] \quad \text{Eq. 2.1}$$

where, I is the PV module output current, I_L is the module photocurrent, I_0 is the saturation current, q is the electron charge constant, γ is PV curve fitting parameter, k is the Boltzmann constant, T_c is the module temperature, V is the PV module voltage, and R_s is the module resistance. The major parameters used in PV panel TRNSYS model are summarized in Table 2-3.

Table 2-3. PV panel model parameter

Parameter	Value	Unit
Open circuit voltage	21.6	V
Short circuit current	6.5	A
Maximum power voltage	17	V
Maximum power current	5.9	A
Temperature coefficients at open circuit voltage	-0.079	--
Temperature coefficients at short circuit current	0.02	--
Nominal operating cell temperature (NOCT)	298	K
Insolation at NOCT	800	W/m^2
Number of modules in series	51	--
Number of modules in parallel	6	--
Module area	0.89	m^2

The overall component model connection in TRNSYS is illustrated in Figure 2-3, where weather condition model reads in weather data, such as outdoor temperature, solar radiation, at each time step from typical meteorological year weather file, and converts them into a desired system of units which generate all the weather variables that PV panel model needs. The weather file used in this module is the same as the one used in Building module. Both of them are the

same TMY3 weather files for Philadelphia. Of course, such TMY3 weather file can be replaced with real weather condition files and for other locations. In this project, the PV panel model can be in one of the following four states: charging battery, powering building, selling power to grid or being dormant. The “Matlab Controller” component is connected with the overall “Operation Module” which passes the overall operation signals into the PV-Battery module, such as “PV charging battery”, “PV powering building”, and “Battery charging building”, etc..

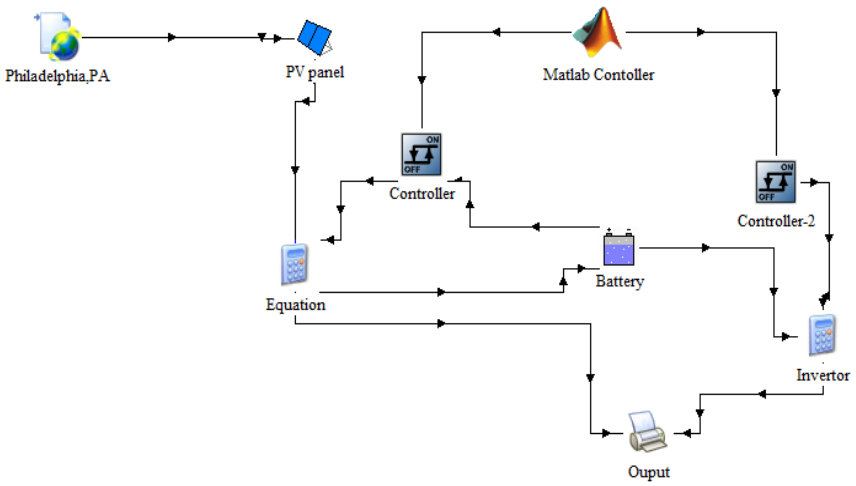


Figure 2-3. PV-Battery system model in TRNSYS

The battery model was developed using Type 47b in TRNSYS, which simulates a lead-acid storage battery in conjunction with PV panel. It specifies how the battery state of charge varies over time, given the rate of charge and discharge. As introduced above, the operation of battery model is also determined by the “Matlab Controller” component. Similar to PV panel power generator, the states of battery can be in only one of the following three states: charging from the PV panel, charging from the power grid, discharging to power the building. The parameters used in this project are illustrated in Table 2-4. The output of this module is the power provided to the buildings in this cluster.

Table 2-4. Battery model parameter

Parameter	Value	Unit
Capacity	1380	Ah
Nominal voltage	12	V
Voltage at no charge	1.9	V
AC/DC inverter efficiency	0.9	--
Maximum round efficiency	0.9	--
Minimum round efficiency	0.7	--

2.3.3 Ice Tank Thermal Storage Module Development

Ice tank thermal energy storage system is another building energy management equipment, which is often used to shave the high electricity demand from cooling load during peak hours associated with real time electricity price. The default ice tank model (object) in EnergyPlus is chosen in this study for its robustness and correctness. There are two different ice tank thermal storage objects in EnergyPlus: simple model and detailed model. The detailed ice storage object in EnergyPlus is used in this project. This detailed model allows user defined charging and discharging curves to model a specific ice storage device more closely [171]. There are three different operation models in this model, namely dormant mode, charging mode and discharging mode. In charging or discharging modes, the state of charge (SOC) of the ice tank can be calculated from charging or discharging rate (u), as shown in Eq. 2.2. During the discharging period, the ice tank storage system provides cooling to meet the cooling demand from the demand side. The TES water flow rate, $\dot{m}_{ice}$, is adjusted based on the load request, $\dot{Q}_{ice}$, and inlet chilled water temperature, T_{inlet} , as Eq. 2.3 [37]:

$$SOC_t = u\Delta t + SOC_{t-\Delta t} \quad \text{Eq. 2.2}$$

$$\dot{m}_{ice} = \frac{\dot{Q}_{ice}}{C_p(T_{inlet} - T_{stp})} \quad \text{Eq. 2.3}$$

Since the default ice tank storage system model in EnergyPlus has to attach to a specific building, which cannot be shared in multiple buildings, a novel ice tank model configuration has been developed in this study. As illustrated in Figure 2-1, the ice tank system was modeled as an individual EnergyPlus model with a dedicated chiller. While with the ice tank model as an individual EnergyPlus model, it is very important to share and exchange "ice cooling" from ice tank to building EnergyPlus models. In the ice tank model, chilled water discharged from the ice storage system is sent to the different buildings separately. The ratio of the chilled water mass flow rate for each building is determined by the Control Module. The cooling load of each building is covered firstly by the chilled water from the shared ice storage tank. And the remaining cooling need is satisfied by the base chillers of each building. A dedicated chiller is used to charge the ice storage system. The charging and discharging schedule is controlled by the Control module through dedicated chiller chilled water outlet temperature setpoints [171]. Key parameters of the ice storage system modeled in this are summarized in Table 2-5. Because it is very difficult to actually pass parameters among different EnergyPlus models, the following schemes are used to mimic the mass and heat transfer between the ice storage system and the buildings described above. A user defined "load profile" model is added in the ice storage tank chilled water loop to represent the chilled water request from the buildings. The amount of chilled water request is determined in the Control Module. Therefore, the ice storage tank will cover the load request in this "load profile", representing the coverage of the building's cooling requests. The overall schematic of the ice storage system is illustrated in Figure 2-4a. To represent the cooling provided by the ice storage system in each building model, a new user defined component, "ice cooling", is created in building EnergyPlus model. The overall function of this "ice cooling" component is to cover some part of the cooling load of each building provided by the ice storage

tank. Then the remaining cooling needs will be covered by the base chillers in the building models. The building chilled water loop configuration is illustrated in Figure 2-4b.

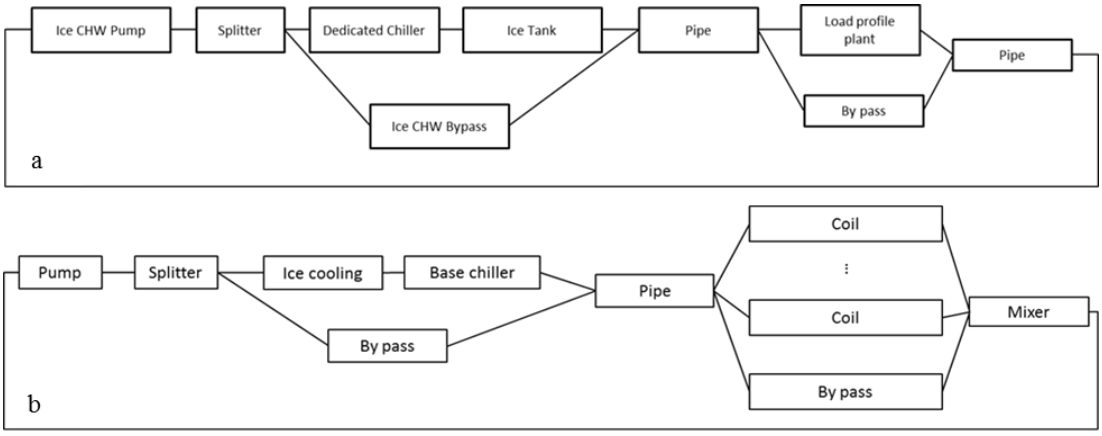


Figure 2-4. Ice tank sharing configuration

Table 2-5. Ice thermal storage tank and dedicated chiller parameter

Parameter	Value	Unit
Ice Tank Capacity	0.2(15.8)	GJ(Tonh)
Tank Loss Coefficient	0.0003	--
Freezing Temperature	0(32)	C(F)
Dedicated Chilled Capacity	7000(2.1E+4)	W(Btu/h)
Dedicated Chilled COP	3.2	--

Similar to a series chiller configuration, the new “ice cooling” component is in a series configuration before the base chilled. The detailed schematic of “ice cooling” is illustrated in Figure 2-5.

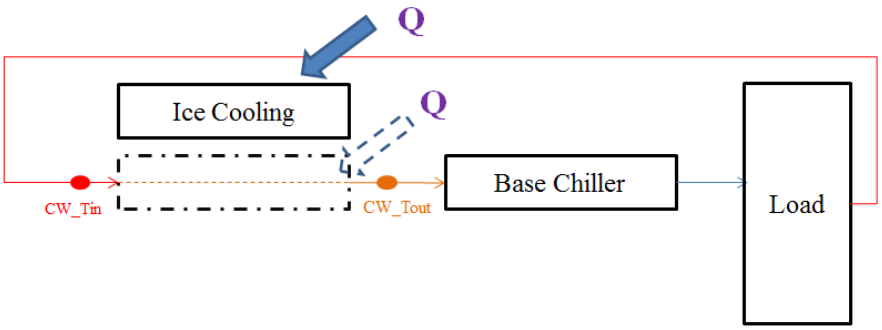


Figure 2-5. “Ice cooling” component schematic

The temperature of returning chilled water will be decreased by the "ice cooling" component and then sent into the base chiller:

$$\mathrm{T_{CWout}} = \mathrm{T_{CWin}} - (\mathrm{Q_{ice}}/(\mathrm{M_{CW}} \cdot \mathrm{CP_{CW}}) \quad \text{Eq. 2.4}$$

where, T_{CWout} is the chilled water temperature at the outlet of the ice cooling component, which is also the chilled water temperature at the inlet of base chiller:

T_{CWin} is chilled water temperature at the inlet of the ice cooling component;

Q_{ice} is the building request cooling from the ice storage tank;

M_{CW} is chilled water mass flow rate in the chilled water loop;

$\mathrm{CP_{CW}}$ *is the specific heat of chilled water.*

2.3.4 Virtual Power Generator Model Development

Since the PV panel and battery system is modeled in TRNSYS, a virtual power generator model is created in the building models (in EnergyPlus) to represent the amount of electricity provided from the PV-Battery module (Figure 2-6). This virtual power generator will be controlled by "TrackSchedule" scheme [171] in to generate the same amount of power as that simulated in the "PV-Battery" module which is used to power building. This power generation value was first simulated and calculated in TRNSYS and then read into MATLAB "Operation Module", and finally sent to the virtual power generator in EnergyPlus through the BCVTB.

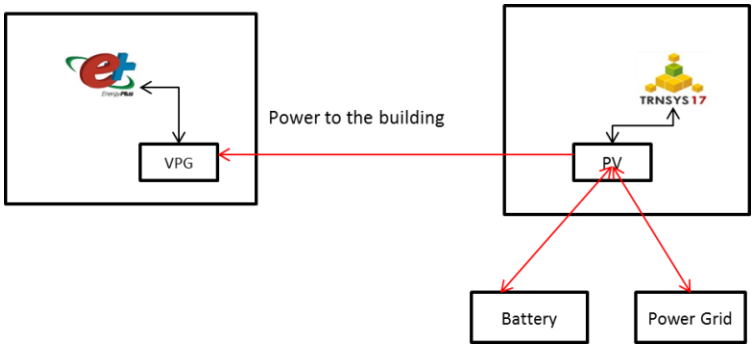


Figure 2-6. Virtual power generator (VPG) configuration

2.3.5 Power Grid Model Development

In this study, a simplified power gird model is also developed to provide the time-of-use electricity price and to calculate the electricity cost and earning based on the amount of electricity buying from and selling to the power grid. This information will be sent to the Control module to assess and determine an optimized control/operation strategy. Figure 2-7 shows the TOU electricity price plan used in this project (SCE, 2008). The following equations are used in this model:

$$C_g = \sum_{i=1}^{m} \sum_{j=1}^{H} (E_{p,i,j} P_{tou,j} - E_{s,i,j} P_{s,j}) \qquad \text{Eq. 2.5}$$

Where, C_g is the net electricity cost ($), m is the number of the buildings in this cluster emulator, H is the building operation time, and all other variables have been introduced in Table 2-1. $E_{p,i,j} P_{tou,j}$ calculates the energy purchasing cost ($), and $E_{s,i,j} P_{s,j}$ calculates energy selling earnings from power grid ($).

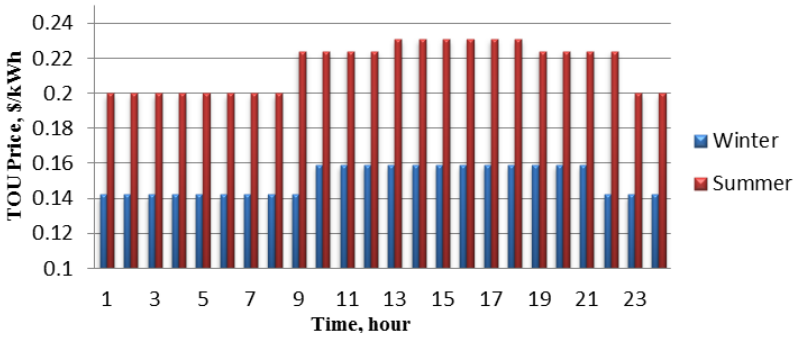


Figure 2-7. Time-of-Use electricity price

All these sub-modules in this building cluster, such as building models, PV-Battery models, and Ice tank storage models are connected to the power grid, where electricity price plays a key role in building operation. It is noticed that a number of public utility commissions and utilities started to use time-of-use (TOU) electricity price or even real-time price. Real-time pricing usually is more focusing on large commercial and industrial customers. For small commercial buildings, TOU price is much more common.

2.3.6 Operation Module Development

The major role of operation Matlab module is to provide predefined operation strategies, which also can work with some optimization models to determine the building cluster operation signals. However, the objective of this emulator is not to look for the building optimal control signals for the building cluster, but to simulate its operation and provide real-world-like operation data to access the operation strategies and to develop on-line estimation models. All the predefined control signals will be put together in a vector to BCVTB server and then sent to different building models, PV panel power generation model, and ice storage tank model at every simulation time step, respectively. The format of this vector is shown in Eq. 2.6:

$$S = [T_{ice}, Q_{iceall}, M_{ice}\ T_{Hi}, T_{Ci}, P_{PVi}, Q_{icei}] \quad \text{Eq. 2.6}$$

where, T_{ice}, Q_{iceall}, and M_{ice} are signals controlling the ice storage tank. T_{ice} is the ice storage tank inlet water temperature setpoint, which is used to control the charging and discharging schedule . Q_{iceall} is the total cooling that the ice storage tank need to provide, which is the summation of all the building request. M_{ice} is the mass flow of chilled water provided. Q_{iceall} and M_{ice} are inputted into the load profile plant component. T_{Hi}, T_{Ci}, P_{PVi}, and Q_{icei} are the control signals to each building, i is the building number. T_{Hi} and T_{Ci} are heating and cooling temperature setpoints. P_{PVi} is the electricity provided to the building from PV power generation. Q_{icei} is the cooling provided to the building from ice storage tank.

The operation of ice storage thermal storage model in EnergyPlus was controlled by the dedicated chiller outlet temperature and the ice storage outlet temperature setpoints. There details on the ice tank operation control in EnergyPlus can be found in EnergyPlus Input-Output Reference [171].

2.4 Emulator Realization in the BCVTB

In this project, all different modules/models are connected with the MATLAB through BCVTB, as shown in Figure 2-8. The “EnergyPlus I” and “EnergyPlus II” simulators are the two building models. The “Ice tank” simulator is the shared ice storage tank EnergyPlus model. The “Matlab” simulator is the Control module and PV-Battery TRNSYS model connector, which will provide operation and control signals and call TRNSYS model every time step. To be more specific of the connection, an external interface is created in these three EnergyPlus models, which is used to connect EnergyPlus and BCVTB. Once EnergyPlus model and BCVTB are connected, the control signals (Input) for Building Module are transferred though this interface from BCVTB. Meanwhile, the control signals from Control module are sent to BCVTB through a shared socket connection from MATLAB. TRNSYS is connected to MATLAB through modeling

calling commands, and simulation results from TRNSYS are imported into MATLAB simultaneously. Control signals for PV-Battery model are injected into their TRNSYS models' "dck" files at each time step. More details about the software connections can be found in Appendix A.

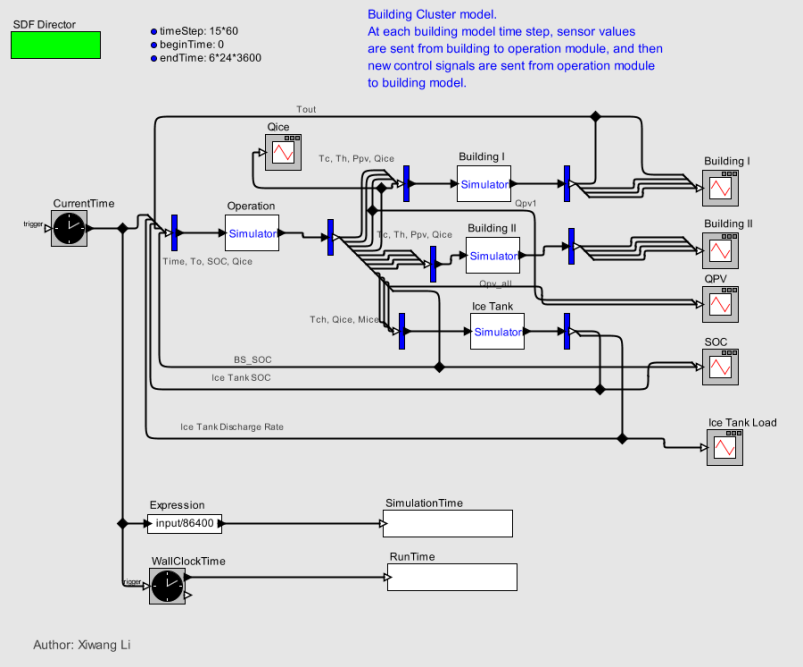


Figure 2-8. Emulator connection in BCVTB

2.5 Emulator Proof-of-Concept Operation Testing

2.5.1 Emulator Testing Scheme

The development and connection of each module in the building cluster emulator have been discussed. In this paper, all the models used in this emulator have been validated in previous studies. Therefore the operation testing will focus on the model connection and data exchanging. The overall testing criteria are summarized in Table 2-6.

Table 2-6. Emulator proof-of-concept testing criteria

Category	Model	Checking Variable	Checking Method
Building Control and operation	EnergyPlus	Tsepc, Tseph, Troom Epv, Ebldg	Compare results of Eplus and Matlab Compare TRNSYS and Eplus
Ice tank operation	EnergyPlus	Tchi, Qload, Qdis, SOC_ice	Tchi for charging and discharging, Q request and provide, SOC_ice
PV-Battery system operation	TRNSYS	Spv, Ppv, Pba, Pgrid, SOC_ba	Control signals for charging and discharging, power generation, power to building, grid, and battery, SOC_ba

Results from different models and different software will be compared and validated. The operation of all these models is checked against all the control signals. For example, ice tank should be in charging state, if the dedicated chiller outlet temperature setpoint is -7^{o}C; and it should be in discharging state, if the temperature setpoint is 90 ^{o}C. In order to run this building cluster emulator, all the control variables need to be predetermined. They can be obtained either from optimization model, or from other predefined operation strategies. It is common knowledge that it is very hard to use pure physics based model to calculate the optimal control strategies, because developing and validating this emulator is expensive and very time consuming, in additionally, the simulation speed is not fast enough to be used in the searching based optimization method. Therefore, in this section, some predefine operation strategies will be applied in this emulator.

Table 2-7. Building and ice tank setting for summer (winter)

Time	Building I		Building II		Ice Tank		
	Heating Setpoint, ºC	Cooling Setpoint, ºC	Heating Setpoint, ºC	Cooling Setpoint, ºC	Dedicated Chiller Setpoint, ºC	Load request Building I, W	Load request Building II, W
0-8	18 (18)	28 (26)	20 (18)	22 (25)	-7 (off)	0 (0)	0 (0)
8-12	20 (22)	22 (24)	24 (24)	26 (26)	90 (off)	13,000 (0)	9,000 (0)
12-18	20 (22)	22 (24)	24 (24)	26 (26)	90 (off)	13,000 (0)	9,000 (0)
18-20	20 (22)	22 (26)	24 (24)	26 (26)	45 (off)	0 (0)	0 (0)
20- 24	18 (18)	28 (26)	18 (18)	28 (25)	-7 (off)	0 (0)	0 (0)

All the temperature setpoints and ice tank load declaration are illustrated in Table 2-7. Two different set of heating and cooling setpoints are applied in the building cluster for summer and winter. Ice tank is charging at night and discharging from 8 am until 6 pm. But this ice tank will not be charged when it is full of charge and cannot discharge when it none of charge. Therefore, the ice tank will discharge chilled water from 8 am until it is deplete, and will be charged by the dedicated chilled from 8 pm until it is full of charge. Between 6 pm and 8 pm, the ice tank will keep dominant. The operation of PV panel and battery system is as shown in Table 2-8. Similar to ice tank thermal energy storage, battery cannot be charged when it's full of charge and will not be able to provide electricity to the buildings. The building cluster test operation data under these operation strategies will be discussed in next section.

Table 2-8. PV-Battery operation setting

Time	PV panel	Battery
0-9	Charging battery	Charged by PV panel
9--12	charging battery first and the rest powering building	Charged by PV panel
12-13	powering building	dormant
13--16	powering building	powering building
16--19	powering building	dormant
19--24	dormant	dormant

2.5.2 Emulator Testing Results

2.5.2.1 Building Cluster Simulation Results

The summer case proof-of-concept study is conducted in August 1st for summer case, and in January 3rd for winter case. The location of this building cluster is Philadelphia, PA, USA. Figure 2-9 and Figure 2-10 are showing the temperature simulation results of these two buildings from EnergyPlus for summer and winter simulations. These two building models use the same weather condition file and internal load schedule, but different temperature setpoints. The purple lines are

building cooling setpoints, black lines are building heating setpoints, red lines are outdoor air temperature, and blue lines are average room temperature. From these two charts, the heating and cooling setpoints are identical to their operation settings in Table 2-7, and the building room temperature are effected by the temperature setpoint. In Figure 2-10, room temperature is out of control from 8 am to 6 pm in the summer case. That because the cooling demand exceeds the capacity of air conditioning system. The room temperature is under control in all other cases.

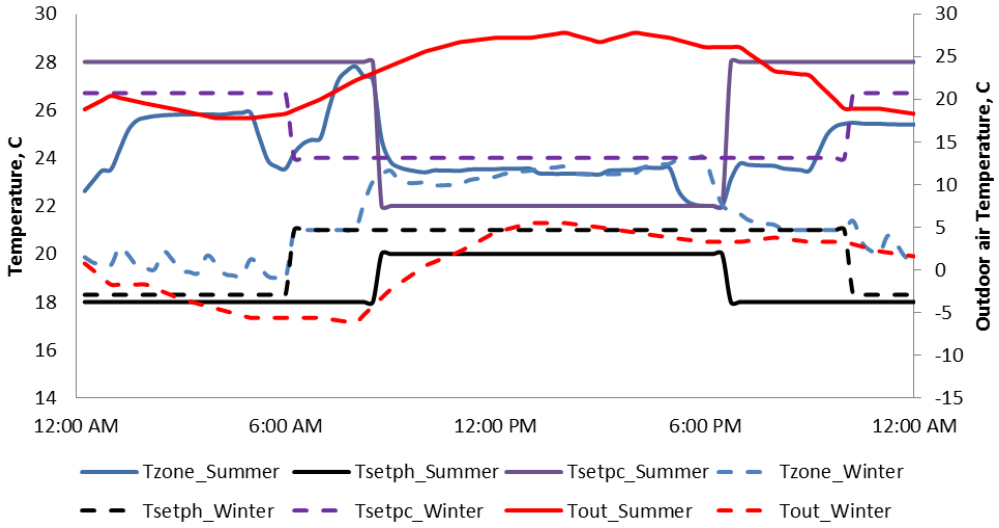


Figure 2-9. Building I temperature simulation results

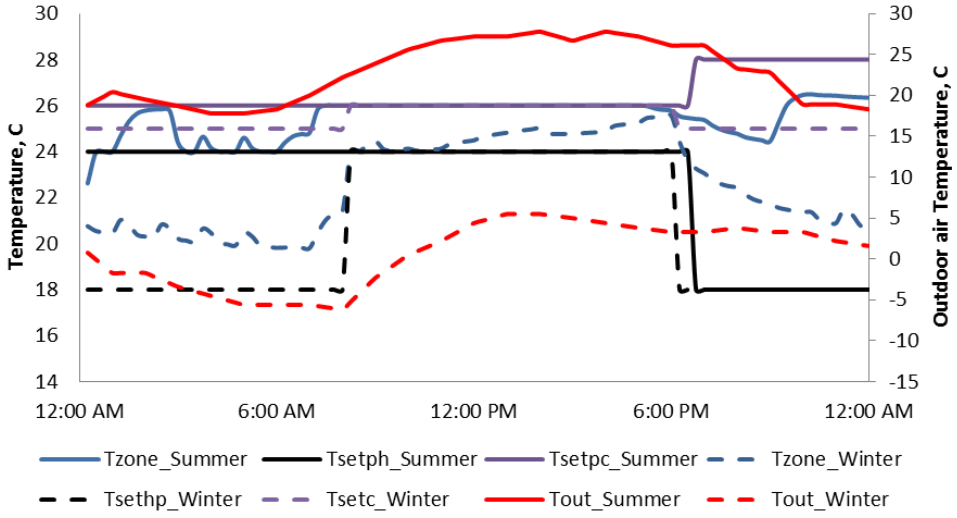


Figure 2-10. Building II temperature simulation results

Similarly, the PV-Battery module simulation results are shown in Figure 2-11. Identical to the operation signals in Table 2-7, PV panel starts to charge battery from easily morning and power the building when battery is full of charge. Starting from 12, PV panel and battery power building together to shave the building peak demand. In the PV-Battery model, 60% of the outputted power is provided to building I, which is the same as the control signals.

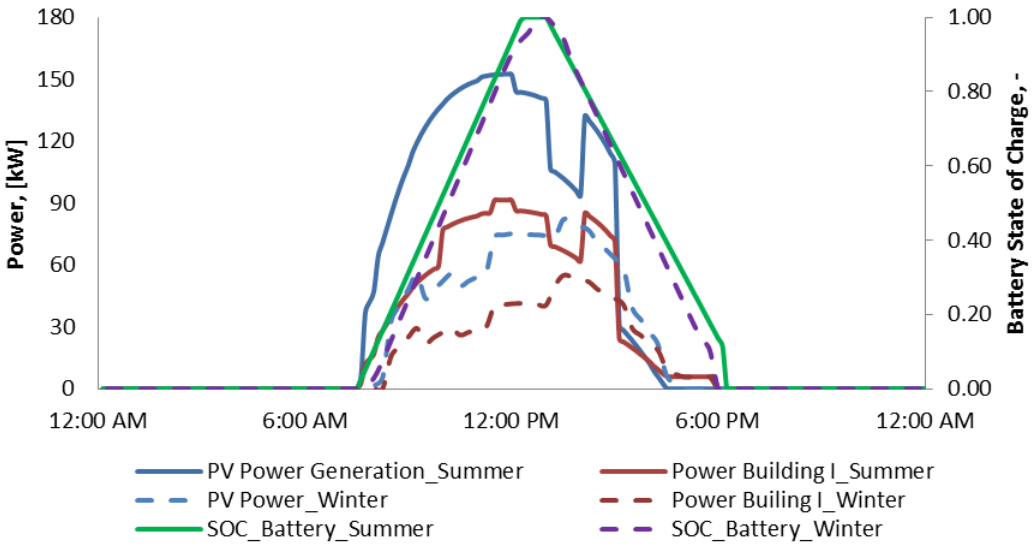


Figure 2-11. PV-Battery operation results

As predetermined in Table 8, the shared ice storage tank is covering part of load for building I and building II from 9 am to 21 pm, until the ice tank is deplete. The simulation results are shown in Figure 2-12. The blue line shows the ice tank storage state of charge (SOC), which stays full until 8 am when it starts to discharge until it is deplete. During the discharge period, the dedicated chiller is also working to meet the cooling requirement from the buildings. The green line is the amount of ice cooling providing to building I, which is 60% of the total discharged rate (red line). This test validates the ice tank share method in section 2.3.3.

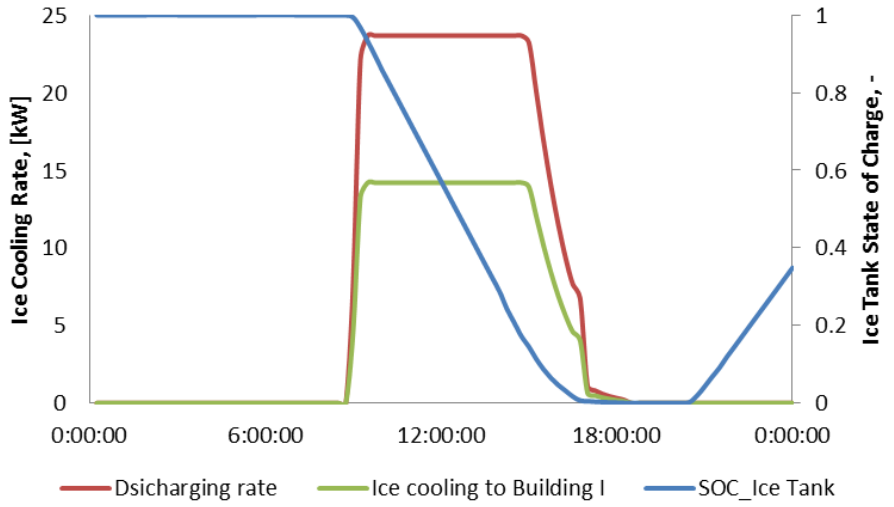


Figure 2-12. Ice tank operation results

2.5.2.2 Building Cluster Energy Consumption and Cost

The energy comsuption and its cost of the building cluster under the testing situation are shown in Figure 2-13 and Figure 2-14. The energy cost doesn't include peak demand charging, which is usually for large comercaial and industral buildings.

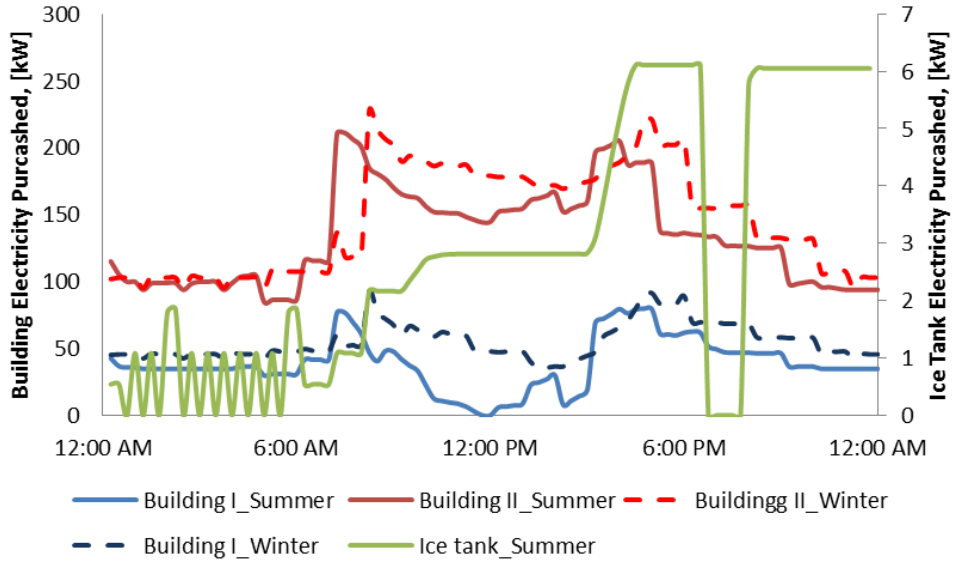


Figure 2-13. Building electricity consumption

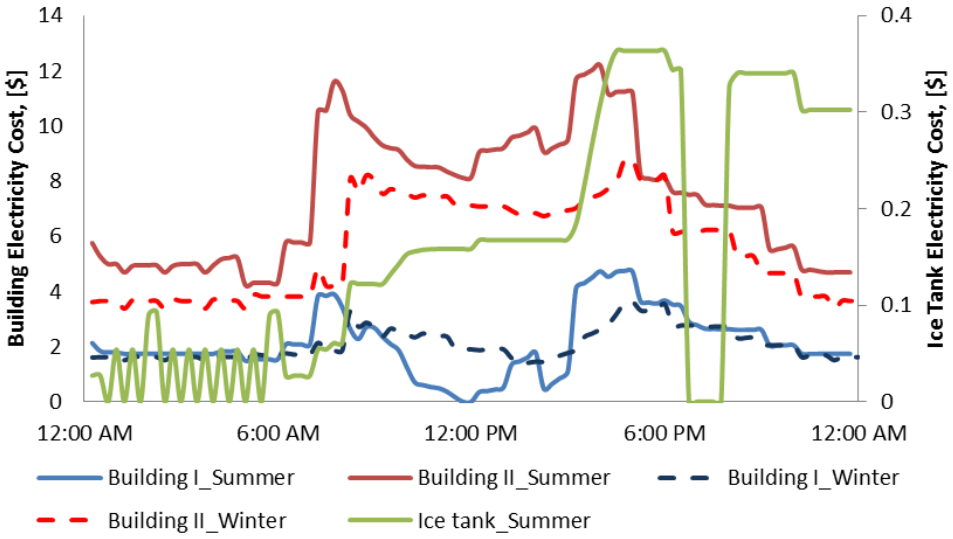


Figure 2-14. Building electricity cost

The energy comsuption at peak hours, from 12 to 4 pm, is reduced in these two buildings, due to the power generation PV panel and battery power discharging, which would save significant energy cost. During off peak hours, ice tank would be charged by the power grid, utilizing the lower electricity price. Building type and thermal mass greatly affect the building energy comsuption and operation. In this study, building II is a larger building comparing with building I, which comsumes more energy, as shown in Table 2-9. The peak demand of building I is around 80 kW which is less than the peak PV-Battery power providing. Therefore, the electricity purchasing is zero at noon for about 30 mintues. From the simulation results, the PV-Battery power provided 886.2 kWh electricity and saved $205.6, without considering the peak demand charging.

Table 2-9. Building cluster electricity consumption summery

	Electricity Consumption, kWh		**Electricity Cost, $**		**Peak Demand, kW**	
	Building I	Building II	Building I	Building II	Building I	Building II
Summer Case	927.5	3214.8	203.1	711.5	79.7	210.9
Winter Case	1325.2	3520.0	291.7	781.7	92.6	227.7

2.6 Emulator Uncertainty Discussion

The uncertainties of the emulator model reside in two different categories: buildings and devices' energy models, and energy prices. Comparing building I and building II, however, building thermal mass as an important building character plays a significant role in building operation and energy savings. The weather conditions, such as outdoor air temperature, solar radiation, also greatly affect the energy consumption and energy generation. On the other hand, energy costs of the cluster during off-peak and on-peak periods show that building operation, especially the operation of energy generation and storage devices, should be updated according to the electricity price changing in order to fully utilize electricity at low rate in off-peak hours.

2.7 Conclusions Future Work

An emulator testbed which is designed to assess building cluster operation strategies and simulate a building cluster as well as energy generation and storage devices is developed in this study. In this building cluster emulator, multiple buildings are connected and are able to share PV panels as energy generation devices, and battery and ice tank as energy storage system. Various simulation environments, including EnergyPlus, TRNSYS, are used to model different parts of this building cluster and are inter-connected through the BCVTB to the MATLAB Model. A proof-of-concept test case is conducted to illustrate the use of this testbed and to verify the data exchange within this emulator.

Future work of this study will focus on developing the connection of this testbed with a real BACnet interface and further with real building control systems. BCVTB and MLE+ provide the interface to connect BACnet of real building automation system. Existing studies, such as [164-166, 172], have succeeded in applying BCVTB or MLE+ to connect energy simulation models with real building automation system.

3. CHAPTER 3 Building Energy Forecasting: System Identification Feasibility Study

3.1 Background and Motivation

Model based control has become a promising solution for building operation optimization and energy saving. High fidelity building energy online forecasting models are essential for model based control and operation. Accuracy and computationally efficiency are two of the most important requirements for building energy models. The building cluster developed in Chapter 2 is physical based white box model, which will be exclusively calibrated and validated. However, developing this emulator is engineering demanding and simulation of this emulator in computational costly. Existing studies in this area have mostly been focusing on reducing computation burden using simplified physics based modeling approach. However, creating even the simplified physics based model is often challenging and time consuming. Pure date-driven statistical models have also been adopted in a lot of studies. Such models, unfortunately, often require long training period and are bounded to building operating conditions.

Therefore, this study proposes a novel methodology to develop building energy estimation models for on-line building control and optimization using a system identification approach. In this chapter, a feasibility study of using a system identification approach for whole building energy forecasting is conducted. Building operational data generated from the cluster emulator developed in Chapter 2 are used in the development and validation process. Frequency response function is implemented in this on-line modeling approach to capture the dynamics of building energy system and forecast the energy consumption with more than 90% accuracy and less than 1 minute computational speed. A systematic analysis of system structure, system order and system excitation determination are also demonstrated.

3.2 Novel Building Energy Estimation Method: System Identification Approach

3.2.1 System Identification Model Selection

As stated above, the objective of this study is to develop an on-line building energy model using system identification method. Model structure plays the most important role in model forecasting accuracy. Privara et al. summarized and compared different system identification approaches for building energy modeling and control [173]. There are two different categories of system identification models, which are time domain models and frequency domain models. There is large number of modeling and identification approaches developed over the recent years, but few of them are suitable for building energy modeling, neither a universal model that can work for different building types and operation schemes. Building energy systems, especially the HVAC systems, are very complicated nonlinear dynamic systems, which are hard to model and forecast.

In order to develop a relative simple system model, frequency response function approach is applied in this study due to its excellent performance in handling system nonlinearity [174]. Fundamentally a frequency response function is a mathematical representation of the relationship between the input and the output of a system in frequency domain, which can simplify the time domain transfer function and still capture the useful information of the system dynamic responses, as demonstrated in Eq. 3.1:

$$H(j\omega) = \frac{Y(j\omega)}{U(j\omega)} = \frac{S_{yu}(j\omega)}{S_{uu}(j\omega)} \quad \text{Eq. 3.1}$$

Where $Y(j\omega)$the Fourier is transform of system output $y(t)$, and $U(j\omega)$ is the Fourier transform of system input $u(t)$. However, better results can be obtained in practice by computing the frequency response function (S_{yu}) as the ratio of cross-spectrum between input and output to the power spectrum of the input (S_{yuu}) [174]. Then by applying the Inverse Fourier Transform, the Impulse Response Functions (IRF) per measurement channel are obtained.

3.2.2 System Input and Output Determination

Besides the model structure, model input and output selection is also crucial to the accuracy of system identification model. Based on the physics theory and the data availability, those variables which have strong effect to the energy consumption and can be easily obtained from measurement or other sources are chosen as inputs. The inputs and outputs of the whole building energy forecasting system identification model and are tableted in Table 3-1.

Table 3-1. Variables of system identification model

Variable	Variable Name	Type
E_c	Building cooling energy (W)	Output
T_{out}	Outdoor air temperature (C)	Input
$T_{zone,\ i}$	Zone *i* temperature (C)	Input
$R_{in,i}$	Equipment /occupancy schedule in zone *i* (-)	Input
Q_{dir}	Direct solar radiation (W/m^2)	Input
Q_{dif}	Diffuse solar radiation (W/m^2)	Input
V_{oa}	Ventilation flow rate (m^3/s)	Input

Solar air temperature ($T_{sol\text{-}air}$) is a variable used to determine the total heat gain through opaque exterior surfaces to calculate cooling load of a building. It is not an output from EnergyPlus model, nor a measurement from weather station, and then a specific calculation model is created, using Eq. 3.2:

$$T_{sol-air} = T_o + \frac{\alpha I - \Delta Q_{ir}}{h_o} \qquad \text{Eq. 3.2}$$

Where, α is absorptivity of an opaque wall; I is the global solar irradiance (W/m^2); ΔQ_{ir} is extra infrared radiation due to difference between the external air temperature and the apparent sky temperature (W/m^2); and h_o is the heat transfer coefficient for radiation (long wave) and convection (W/m^2K) .

Direct solar radiation and diffuse solar radiation are used to estimate the building heat gain due to the solar transmission through windows. They can be either obtained from weather forecasting information or calculated from global solar irradiance. All of these input variables are categorized into two groups: unexcited and excited inputs. The zone temperature, and equipment and occupancy schedule ratio are excited inputs, which will be excited during a model training period, and the other variables are unexcited inputs, which values are changed naturally. Using the simulation test bed, the zone temperature is excited by changing the zone temperature setpoints. And the equipment and occupancy schedules are excited by updating their on/off schedules in the simulation emulator testbed.

3.3 System Excitation Signal Generation: Experiment Design

In order to develop the spectrum density model and improve the accuracy of system identification model, exciting signals were generated and injected into the system during the training period. These exciting signals include zone temperature setpoints, internal equipment and occupancy schedules. Therefore, certain constraints added into our system excitation process, for example, the boundary of temperature setpoints, the minimum temperature setpoint or equipment schedule updating time span, and so on. Different system excitation strategies, such as Pseudo-Random Binary signal, Pseudo-Radom Sequences, Multi-sine signal, have been discussed and applied in on-linear process systems in different areas [150, 158, 159]. However, there are few publications on building energy system excitation found in existing literature. Pseudo-random Binary Sequence (PRBS) excitation signals (Eq. 3.3) for building temperature setpoints was generated and applied in [76].

$$T_{sp,i}(k)\begin{cases}21, excited\ zone, PRBS = 0\\25, excited\ zone, PRBS = 1\\25, \quad non-excited\ zones\end{cases} \qquad \text{Eq. 3.3}$$

3.3.1 Excitation Signal Generation Function Selection

Sum of sinusoids (SINE) model is used to generate the exciting signals (Eq. 3.4), because sinusoids signals are versatile periodic and can adjust signal shape and character of the power spectrum by adjusting their parameters.

$$U_{\tau+1} = U_{\tau} + \sqrt{2a_{\tau}} sin(\omega_{\tau} tT + \varphi_{\tau}) \quad \text{Eq. 3.4}$$

Where $U_{\tau+1}$ is the excitation signal; $\sqrt{2a_{\tau}}$ is a magnitude scale parameter from 0 to 1; ω is periodic frequency parameter from 0 to 2π; T is the sampling time, and φ is the phase lag parameter from 0 to 2π, which do not affect the signal spectrum. Lowering T will result in a higher frequency bandwidth [158]. Another benefit of using sum of sinusoids input signals is that they enable the user to directly specify the shape and character of the power spectrum. The guidelines of excitation signal generation function parameter determining from Rivera et al. [175] is applied to ensure the signals contains necessary frequency information:

$$\frac{\mathbf{1}}{\boldsymbol{\beta_s \tau_{dom}^H}} \leq \boldsymbol{\omega_{\tau}} \leq \frac{\boldsymbol{\alpha_s}}{\boldsymbol{\tau_{dom}^L}} \quad \text{Eq. 3.5}$$

where, τ_{dom}^{H} and τ_{dom}^{L} correspond to the high and low estimates of the dominant time constant of the system (denote the slowest and the fastest systems time constants) [158]. α_s and β_s are user-decisions on high and low frequency content based on identification requirement. Typically, α_s is 2 and β_s is 3, corresponding to 95% of settling time [175]. In additional to function magnitude scale parameter, periodic frequency parameter and phase lag parameter, harmonic suppression is another important parameter which can decompose the output signal to obtain a more accurate estimate of the linear and nonlinear components.

3.3.2 Excitation Function Parameter Specification and Data Generation

The procedure of excitation signals design is summarized in Figure 3-1. The response time constant of a dynamic system is a measure of how quickly the system responds to an input change. It is usually determined by experiments. For example, the impulse response of a dynamic system can be expressed as:

$$x(t) = (\alpha/T)e^{-t/T} \qquad \text{Eq. 3.6}$$

where, T is the response time constant, α is a state parameter. The response time for the system output, x(t), to reach 95% of its final steady state value after an input change, is defined as $T_{.95}$For the building in this studied project, $\tau_{dom}^{H} = 360$ minutes, and $\tau_{dom}^{L} = 30$ minutes; α_s determines the high frequency content in the excitation signal and represents the response speed. β_s specifies low frequency information corresponding to the system settling time. α_s and β_s are chosen to be 2 and 3 for the 95% of the system settling time, respectively. For the temperature setpoint excitation signals in this project, $T_{max} = 32^oC\ (90^oF)$ and $T_{min} = 10^oC\ (50^oF)$; while for the schedule ratio excitation signals, $R_{max} = 1$ and $R_{min} = 0$.

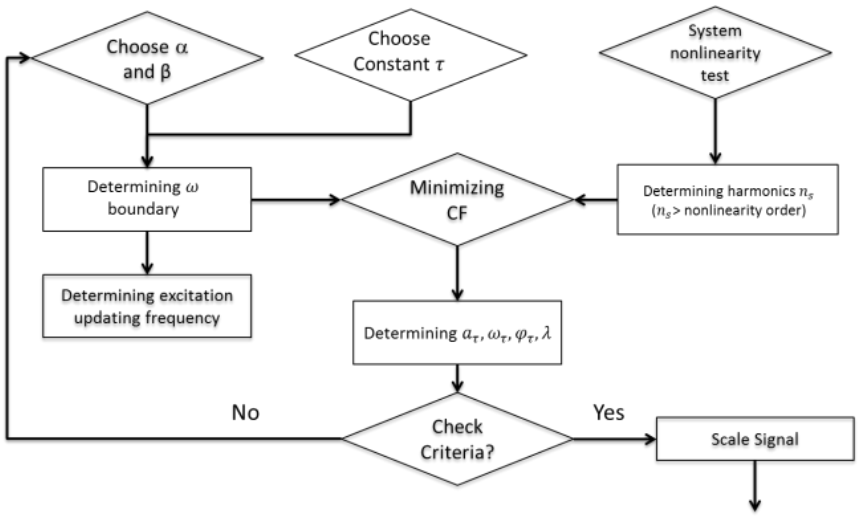


Figure 3-1. Excitation signal generation procedure

3.3.3 Excitation Signal Injection Interval and Data Sampling Length Determination

According to Braun's guidelines, the excitation injection frequency cannot be too high. The injection time step should be larger than the system response time. The building response time is related to building thermal mass, which can only be found through experiment test. The building studied in this project is a small light building. It system response time is relatively short. A parametric experiment test has been conducted to find out the best injection frequency. However, once the system excitation frequency changed, the building dynamics would change, and the power density model sapling length should also be updated accordingly. Table 3-2 presents the excitation frequency and sampling length testing results. Finally, 30 minutes with 6 hours and 30 minutes with 4 hours have been chosen for excitation frequency and sampling length for core zone model and perimeter zone model, respectively.

Table 3-2. Excitation frequency and sampling length testing summary

R^2	**Excitation Frequency**	**Sampling Length**	**Excitation Frequency**	**Sampling Length**	**Excitation Frequency**	**Sampling Length**
	15 Min	3 Hour	30 Min	4 Hour	60 Min	6 Hour
Training Period	94.65%		93.70%		92.76%	
Forecasting Period	97.30%		95.30%		91.91%	

3.3.4 Excitation Signal Evaluation

A primary goal of the system excitation is to lead to produce enough data which can discriminate between any two different models in the data set. The requirement for the informative experiments for open loop operation means that the input should be persistently exciting, and contains sufficiently many distinct frequencies. The covariance matrix is typically inversely proportional to the input power. The desired property of the waveform is defined in

terms of the crest factor (C_r), which shows the ratio of peak values to the average value of the signal (u). The crest factor is defined as:

$$C_r^2 = \frac{max\ u^2(t)}{\lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} u^2(t)} \quad \text{Eq. 3.7}$$

A good signal waveform is consequently one that has a small crest factor. Therefore, a constrained minimum crest factor optimization problem is created to determine the parameters in excitation signal generation function:

$$J(a_\tau, \omega_\tau, \varphi_\tau) = \min_{a_\tau, \omega_\tau, \varphi_\tau} C_r, \tau = 1, 2, \cdots N \quad \text{Eq. 3.8}$$

subject to Eq. 3.9 and other hard boundaries (Eq. 3.10) for building temperature setpoints and equipment schedule:

$$10℃ = U_{T,\tau}^{min} < U_{T,\tau} < U_{T,\tau}^{max} = 32\ ℃ \quad \text{Eq. 3.9}$$

$$0 = U_{E,\tau}^{min} < U_{E,\tau} < U_{E,\tau}^{max} = 1 \quad \text{Eq. 3.10}$$

Following the procedure discussed above, the exciting signals for temperature setpoints and equipment schedules used in this study are generated and shown in Figure 3-2, where all the excitation signals are injected into the building model every 30 minutes.

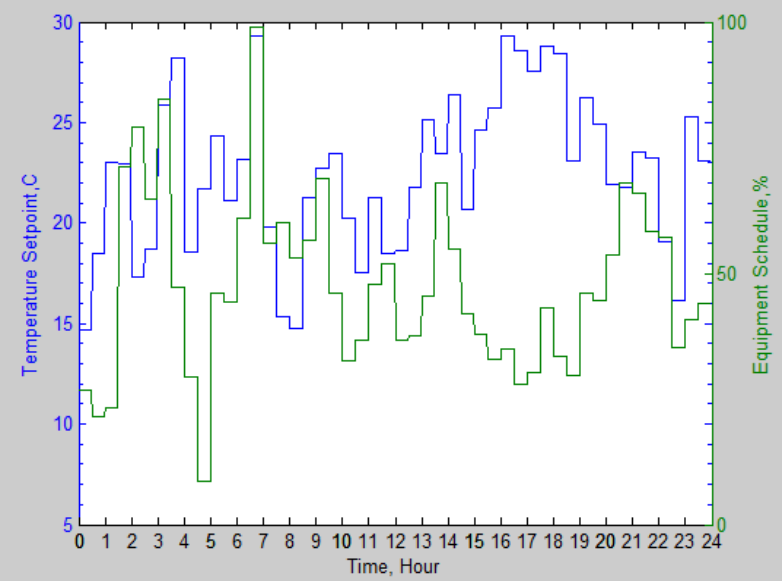


Figure 3-2. Building operation excitement

3.3.5 Training and Validation Data Generation

The system excitation signals discussed in section 2.3 was modeled and generated in Matlab. In order to apply these excitements into EnergyPlus model, BCVTB is used to exchange data between Matlab and EnergyPlus. Here BCVTB plays a master role in data exchange between Matlab and EnergyPlus through run-time coupling, as shown in Figure 3-3. During the entire study, typical meteorological year (TMY) weather data for Philadelphia is used. During the training and validation period, excited and unexcited building control signals will be sent to EnergyPlus model following the procedure in Figure 3-1, respectively. Simulation results and control signals will be sent back and stored in Matlab for system identification model training and validation. The length of training time is changeable according to the forecasting accuracy requirement.

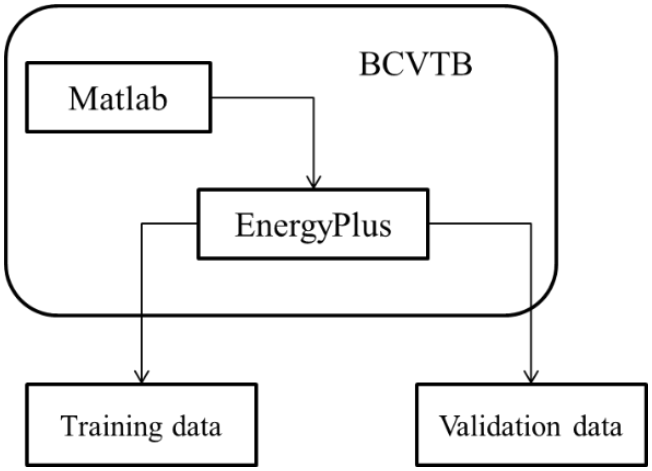


Figure 3-3. Building operation data for on-line model training and validation

3.4 Building Energy On-line Forecasting Model Development

Based on the training data generated from the excited system, a system identification model based on spectral density model for frequency response function is then developed. Figure 3-4 shows the model development process from building operation data. In this figure, *U* is training inputs, *h* is a reference signal to analyze the input data, *Y* is training outputs data, *PSD* is power spectral density model for inputs data and the reference signal, and in this study *h* is Welch spectrum object, *CPSD* is cross power spectral density model for input and output. *Suu* is the result of *PSD*, S_{yu} is the result of *CPSD*, S_{uu} and S_{yu} estimate the correlation between input and output. *G(z)* is the transfer function in frequency domain, which can be transferred to time domain transfer function *G(t)* using inverse Fourier function transformation, and $\hat{y}$ is the output estimation. *G(t)* will be saved as a set of Markov parameters to capture the relationship between each input and output variables.

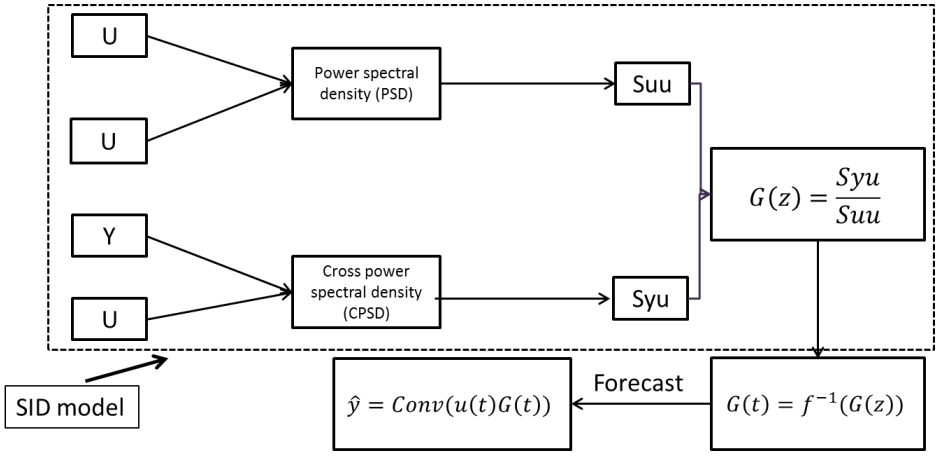


Figure 3-4. System identification model development procedure

Power spectral density (*PSD* and *CPSD*) describes how the power of a signal or time series is distributed over the frequency spectrum, which is a property of the system signal and very useful in frequency domain system identification [174]:

$$S_{uu}(k) = \frac{1}{l}\sum_{\tau=1}^{l-1} R_{uu}(\tau)e^{-j\frac{2\pi k\tau}{l}} \quad \text{Eq. 3.11}$$

$$S_{yu}(k) = \frac{1}{l}\sum_{\tau=1}^{l-1} R_{yu}(\tau)e^{-j\frac{2\pi k\tau}{l}} \quad \text{Eq. 3.12}$$

Where, R_{uu} is the auto-correlation between the inputs and R_{yu} is the cross-correlation between input and output, and l is the length of the sampling data. l is a very important parameter which affect the estimation accuracy and speed, because within one data sample, the power density is calculated simultaneously.

$$R_{uu}(\tau) = \frac{1}{l}\sum_{i=1}^{l-1} u(i)\,u^T(i-\tau) \quad \text{Eq. 3.13}$$

$$R_{yu}(\tau) = \frac{1}{l}\sum_{\tau=0}^{\infty} y(k)u^T(i-\tau) \quad \text{Eq. 3.14}$$

Figure 3-5 shows the procedure of using system identification model (*SID* model hereafter) to develop the on-line building energy model. During the training process, the excited building control signals will be generated in Matlab according to the exciting scheme discussed above and sent to EnergyPlus through BCVTB. The EnergyPlus simulation results of the excited system will be used to train the *SID* model, and calculate its Markov parameters for each input in transfer function. On the other hand, the EnergyPlus simulation results of the unexcited control signals will be used to validate the system identification model.

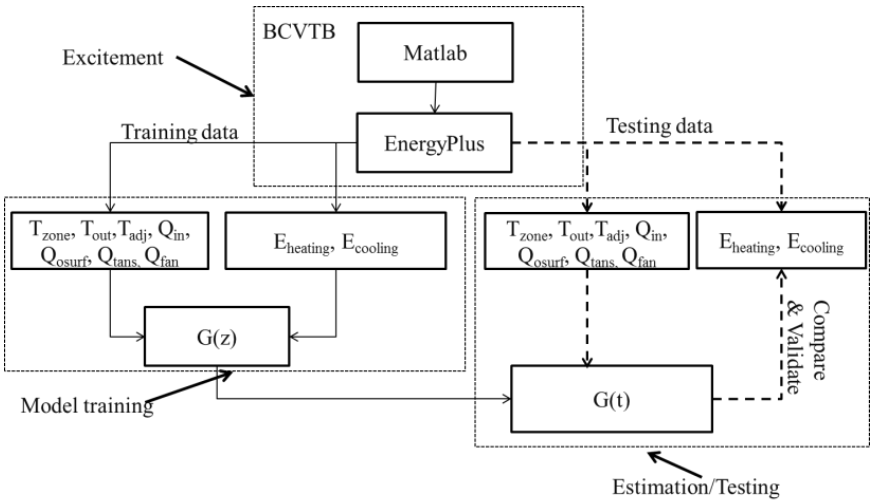


Figure 3-5. On-line building energy model development procedure

3.5 System Identification On-line Forecasting Results

The training period for building heating and cooling energy estimation models in winter is from Jan. 1st to Jan. 10th, and the forecasting period for them is from Jan. 11th to Jan. 13th. While the training period for cooling heating and cooling energy estimation models in summer is July 1st to 10th, and the forecasting period is July 11th to 13th. During training period, all the system exciting signals are applied into the building model, but regular control signals are used in the building during the forecasting/testing period (Figure 3-6).

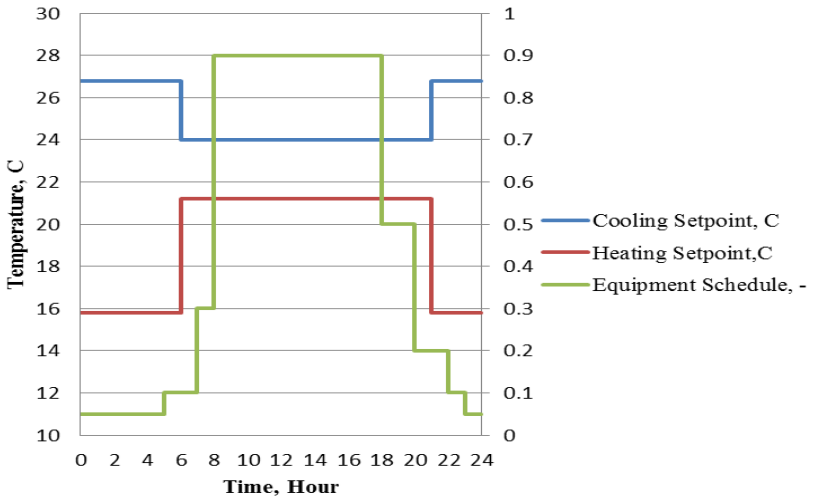


Figure 3-6. Temperature setpoints and equipment schedule during forecasting period

3.5.1 Studying Building Description

The small-size commercial building studied in this project is a single story office building (Figure 3-7), which has six zones, five conditioned zones and an unconditioned attic zone, and the total floor area is 510 m^2. The window-to-wall ratio of this building's facades is approximately 21.2%, and the windows are equally distributed. The overall U-factor of these single pane windows is 3.4 W/m^2K and the solar heat gain factor is 0.36. The solar absorptivity, transmissivity and reflectivity are 0.06, 0.69 and 0.24, respectively. The roof insulation has an R-value of 15. The roof is covered in an asphalt membrane, with a solar absorptivity value of 0.9. The overall U-factor for the walls is 0.68 W/m^2K. The building location is selected as in Philadelphia, PA, USA for this study.

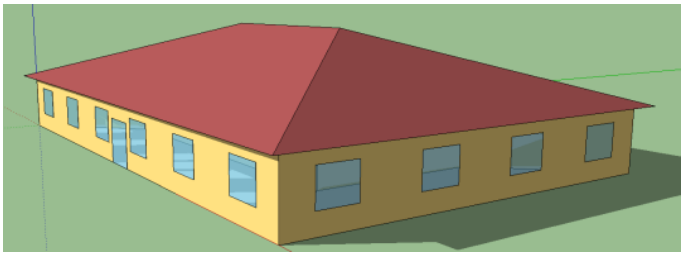


Figure 3-7. Small commercial building view

The HVAC systems used in this building are constant-air-volume (CAV) air handling units (AHUs) with direct expansion (DX) coils. The coefficient of performance (COP) of the cooling system is 3. Heating is provided by electricity with an efficiency of 0.95. The baseline model internal load inputs are summarized in Table 3-3.

Table 3-3. Building model baseline heat gains

Variable	Value
Occupant Density	0.005 person/square foot
Ventilation Requirement	26.5 CFM/person
Lighting Power Density	1.8 watts/square foot
Interior Small Plug Loads	1.0 watts/square foot
Elevator Consumption	32,000 watts
Exterior Lighting	18,000 watts
Envelope Infiltration Rate	0.223 CFM/square foot

3.5.2 Energy Estimation Model Performance Evaluation Index

Model forecasting accuracy when compared with EnergyPlus results and speed are two most important indices in this study.

Coefficient of determination, R^2 (Eq. 3.15) is used to measure the forecasting data accuracy [176].

$$R^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(\hat{x}_i - \bar{\hat{x}})}{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(\hat{x}_i - \bar{\hat{x}})} \quad \text{Eq. 3.15}$$

Where x_i is the energy consumption from EnergyPlus, $\hat{x}_i$ is the energy consumption forecasting data, $\bar{x}$ and $\bar{\hat{x}}$ are their average.

Integrated root-mean-square error (Eq. 3.16) is also used to evaluate the model performance.

$$N_{rms} = \frac{100}{x_m}\sqrt{\frac{\sum_{j=1}^{N}\left(x_j - \hat{x}_j\right)^2}{N-1}} \quad \text{Eq. 3.16}$$

Where x_j is the energy consumption from EnergyPlus, $\hat{x}_j$ is the energy consumption forecasting result, x_m is the mean value of energy consumption from EnergyPlus, N is the total number of time step.

3.5.3 Building Summer Cooling Energy Results

Using the system identification approach developed in pervious chapters, the energy forecasting results are discussed in this section.

Building Cooling Energy Forecasting

Since the building studied in this section is a one-floor building, one overall model for the whole building is developed for the heating and cooling energy estimation. One very important variable was added into energy estimation model, which is supply fan heat into the air stream. The building studied in this project is a light mass building, whose system response time is relatively short. A parametric experiment test has been conducted to find out the best injection frequency. Finally, 30 minutes with 6 hours has been chosen for excitation frequency and sampling length, respectively. The time step for state updating is 15 minutes in the system identification model. Figure 3-8 shows the Markov parameters for each input variable, using the excitation and response time-history data during the sampling window. All the Markov parameters are close to 0 at the end of the sampling period, which means the sampling length is longer enough to capture the influence on the output from all the inputs. These ten series of Markov parameters is then used in the state space model formation and Kalman filter implantation.

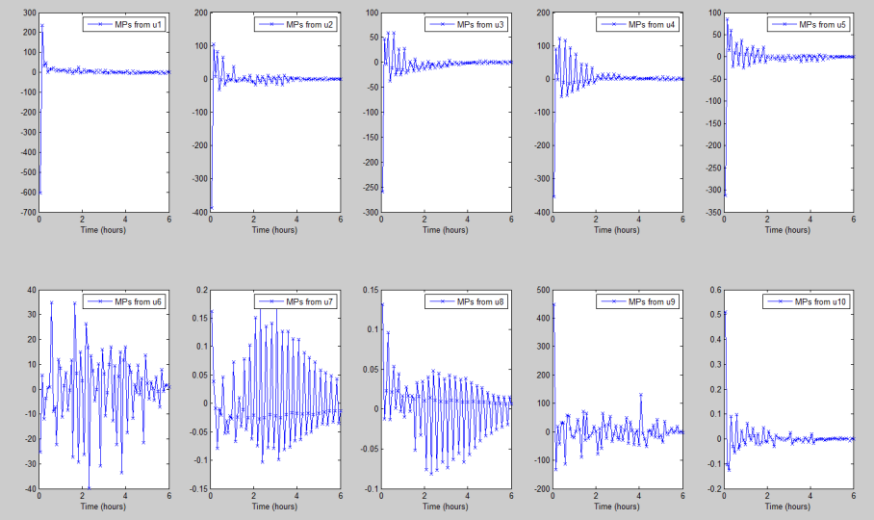
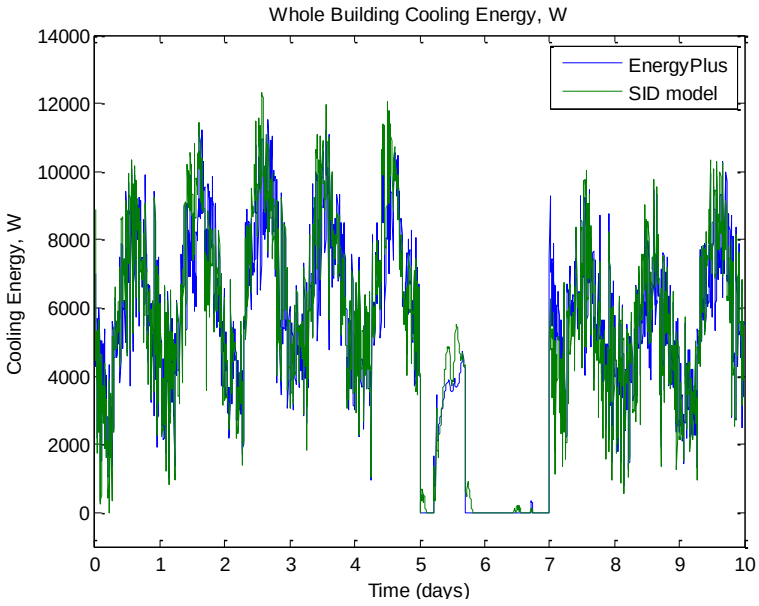


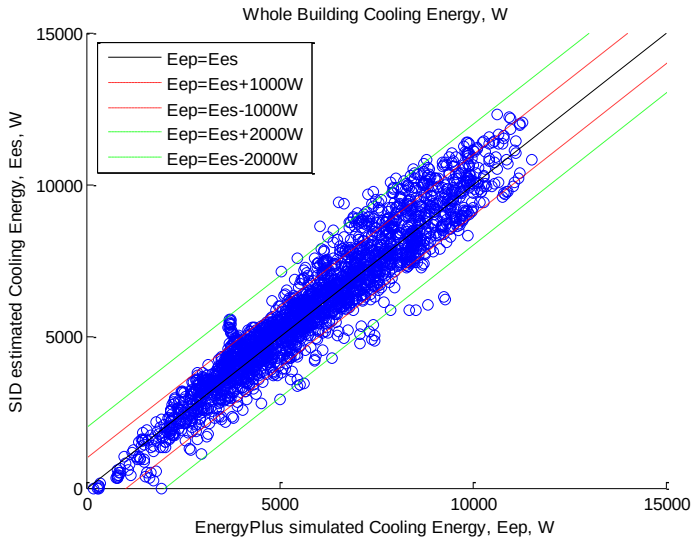
Figure 3-8. Markov parameters from system identification model

The system identification model developed based on the July training data is then used to forecast the whole building cooling energy consumption in the same training period. As shown in Figure 3-9, the model is able to capture the overall trend of building cooling energy from EnergyPlus simulation test bed. The R^2 is 0.94 for the entire 10 days of training data.

The system identification model is then used to forecast whole building cooling energy consumption for three days (July 11 to 13). The results together with simulated results from the test bed are illustrated in Figure 3-10a-b. They illustrate the comparison of EnergyPlus simulated whole building energy consumption (*Ees)* and System identification model (*SID*) forecasted while building energy consumption (*Eep*). Due to the underestimation of direct solar radiation related cooling energy consumption in the afternoons, when cooling load is high, discrepancy between *Ees* and *Eep* exists in the afternoons. Even with this underestimation, the overall forecasting accuracy is still acceptable (Table 3-4). As Table 3-4 shows, the forecasting period accuracy (R^2=0.955) is higher that of training period (R^2=0.944). That is because building energy system has much higher dynamics during the training period when excitation signals are applied.



a) Building energy training estimation results

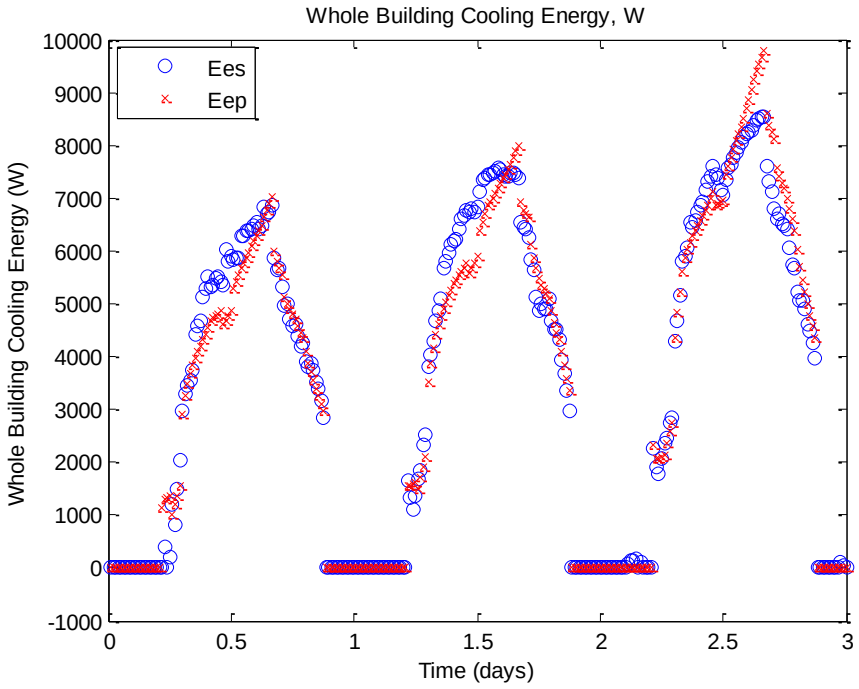


b) Building training error analysis

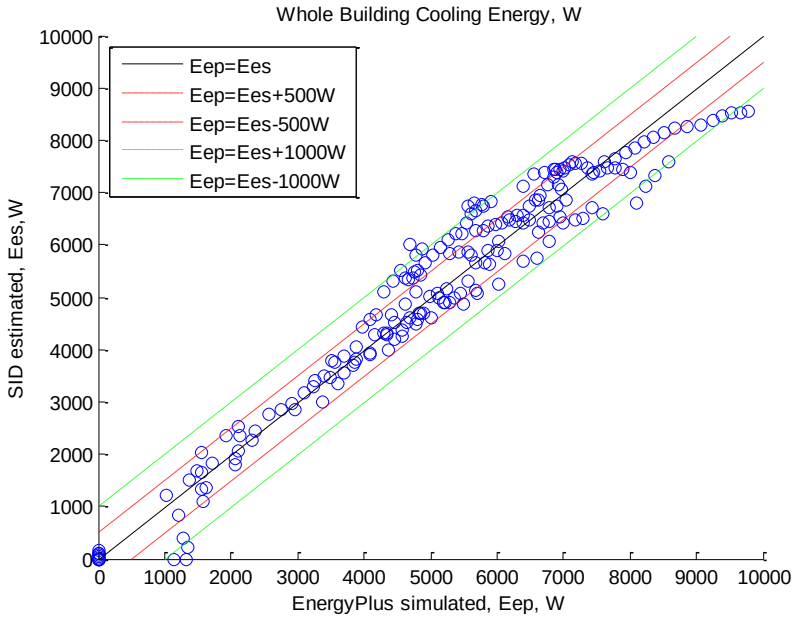
Figure 3-9. Building energy training results

Table 3-4. System identification model cooling energy results from July 11-13

System Identification Model	Accuracy (R^2/N_{rms})		Speed (S)	
	Training Period	Forecasting Period	Training Period	Forecasting Period
Whole Building	0.944/14.7%	0.955/11.03%	47.33	0.0089



a) Building energy forecasting results

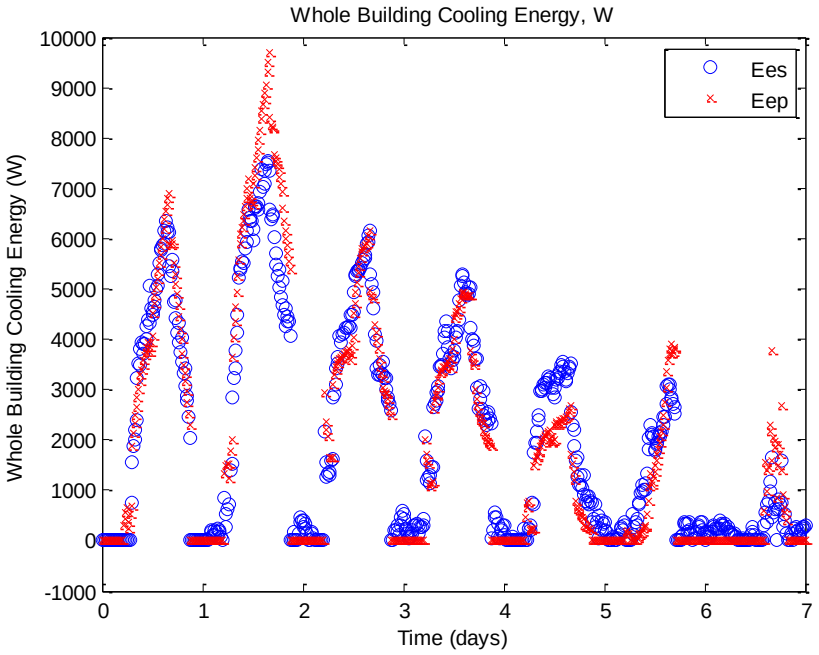


b) Building forecasting error analysis State Space model Estimation Results

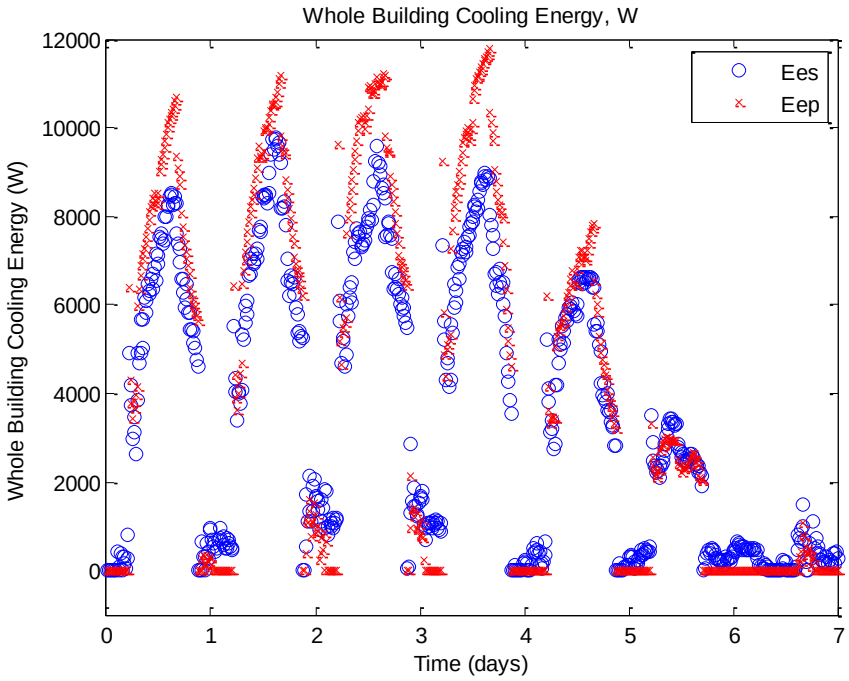
Figure 3-10. System identification model results

3.6 System Identification Model Robustness Test

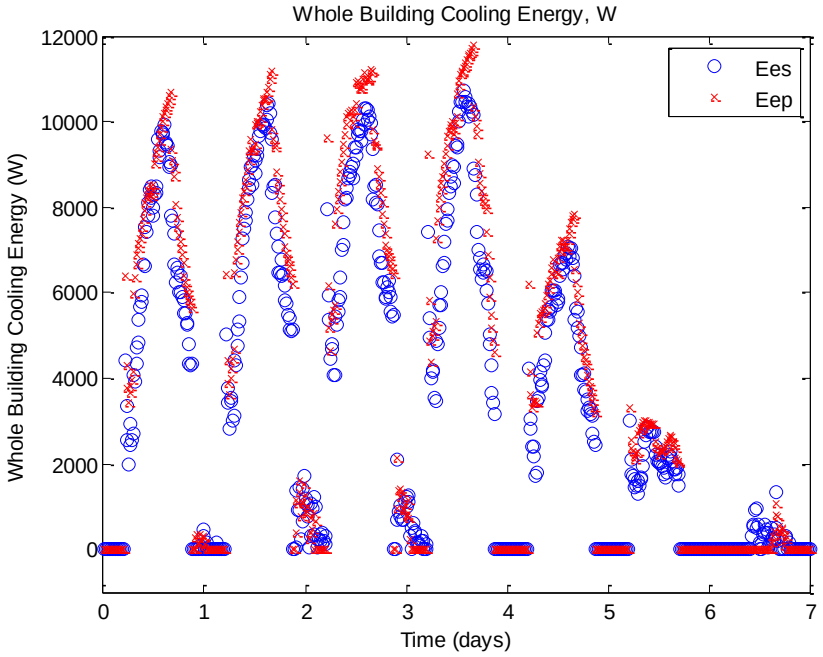
In this section four different case studies are conducted to analyze the model robustness. Since the state space model and the Kalman filter on-line model are based on the Markov parameters calculated in the system identification process, just the system identification model is analyzed and compared in this robustness analysis. The four different case studies are with different training and forecasting periods. The four plots in Figure 3-11 show the forecasting results of these cases, and Table 3-5 summarizes their accuracy.



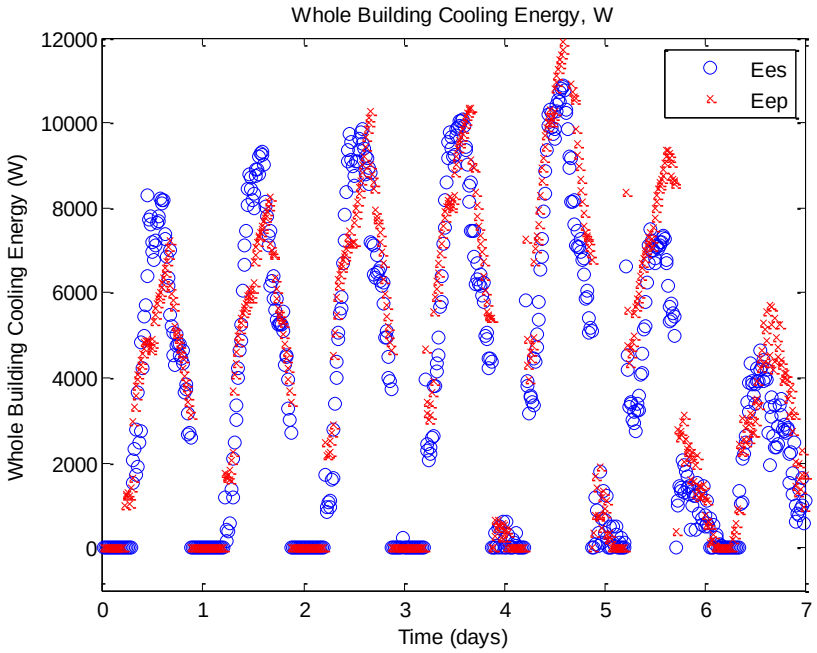
a)Training period July 01- July 10 for July 21- July 27 forecasting



b) Training period July 01- July 10 for Aug. 11- Aug. 17 forecasting;



c) Training period Aug.01-Aug.10 for Aug. 11- Aug. 17 forecasting;



d) Training period Aug.01-Aug.10 for July. 11-July. 17

Figure 3-11. System identification model robustness results

Comparing Figure 3-10 and Figure 3-11, the same model trained from July 01 to July 10 achieved the highest accuracy when it is used to forecast the energy consumption during July 11 to July 13, while it achieved the lowest accuracy for Aug. 11 to Aug. 17. This is reasonable because the weather conditions in Aug. 11 to Aug. 17 are quite different to those in July 01- July 10, while those in July 11 to July 13 are very close to the training conditions.

Table 3-5. Model forecasting robustness testing summery

Training Period	**Forecasting Period**	**Accuracy, R^2**	**Result**
July 01- July 10	July 11- July 13	0.95	Figure 3-10
	July 21- July 27	0.92	Figure 3-11a
	Aug. 11- Aug. 17	0.89	Figure 3-11b
Aug.01-Aug.10	Aug. 11- Aug. 17	0.95	Figure 3-11c
	July. 11-July. 17	0.88	Figure 3-11d

Similarly, the model trained from Aug.01 to Aug.10 (Figure 3-11c) achieved better accuracy then that trained from July 01 to July 10 (Figure 3-11b) when they forecast the building energy consumption in Aug. 11- Aug. 17. Above all, training period weather condition plays a critical role to the model forecasting accuracy. Model forecasting accuracy is higher when the forecasting period condition is closer to the training period conditions.

3.7 System Identification Model Application in Medium Office Building

3.7.1 Medium Office Building Description

The system identification modeling approach developed in this study has been utilized in another medium office building, in order to test its accuracy and robustness. Figure 3-12 illustrates the medium building view in EnergyPlus, which is a 53,628 ft^2, three-story, 15-zone building [177]. The window-to-wall ratio of this building's facades is approximately 33.0%, and the windows are also equally distributed. The thermal properties of the windows of this building are the same as those in the previous small commercial building.

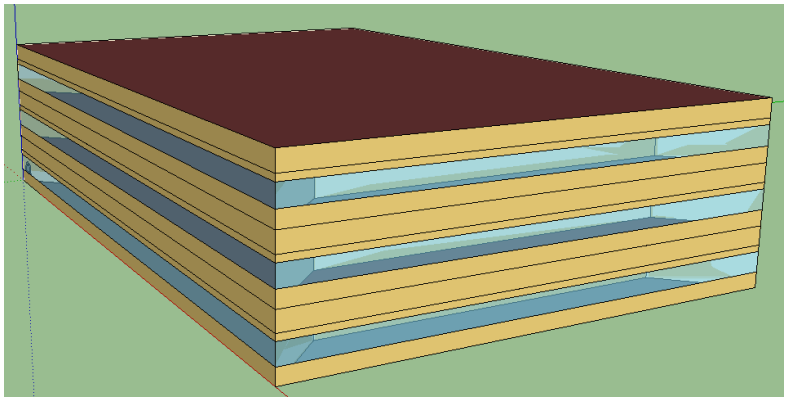


Figure 3-12. Medium commercial building view in EnergyPlus

The mechanical configurations of this building are tabulated in Table 3-6. Different to the small commercial building, this building is using variable-air-volume (VAV) AHU. Primary

cooling is also provided by electricity through DX coils. The coefficients of performance (COP) of these cooling coils are 3. Primary heating is provided by a natural gas (NG) boiler, which have 70% annual fuel efficiency utilization (AFUE). Zone reheat is provided by electric resistance heating.

Table 3-6. Mechanical systems of medium office

	Specification
System	3 VAV, AHUs
Main Cool Coil	DX, COP 3
Main Heat Coil	NG Furnace
Zone Reheat	Electric
Heat Plant	Central Boiler
Heat Efficiency	70% AFUE

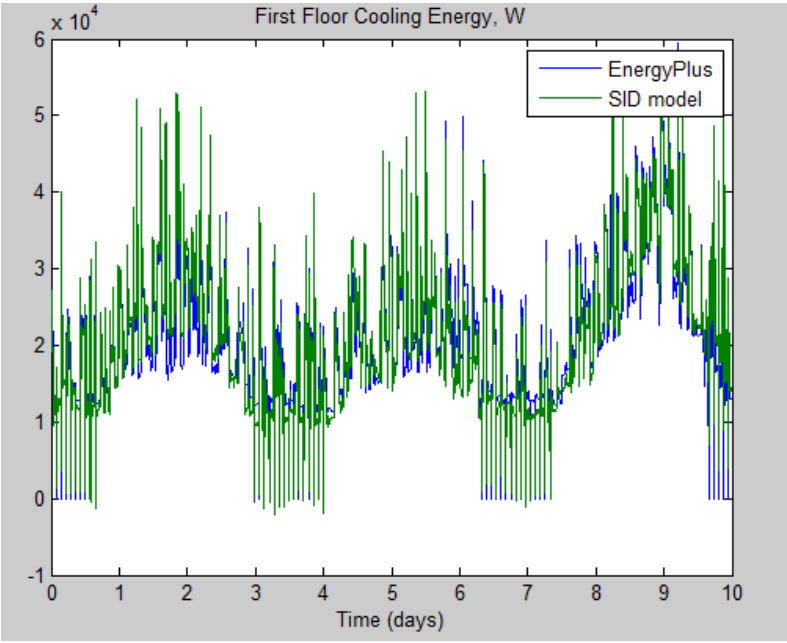
The system identification model developed and validated in previous sections have been applied in this building with inputs adjusted based on the system configurations. Since there are 3 floors in this building, three different models are developed for each floor which consists of five zones: core zone, east perimeter zone, south perimeter zone, west perimeter zone and north perimeter zone. Same system excitation strategy was used in this building for building temperature setpoints and indoor equipment schedules. The inputs and outputs of each model are also similar to the small building model. But in each model, temperature of each room and each adjacent room at each floor are included.

Table 3-7. System identification model forecasting accuracy and speed for medium building

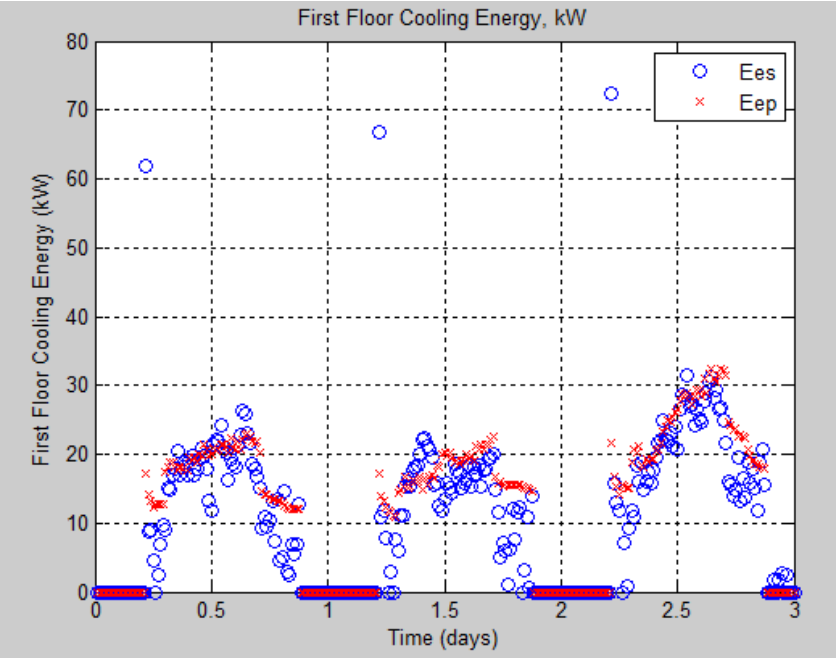
R^2	**Accuracy (R^2)**		**Speed (S)**	
	Training	**Forecasting**	**Training**	**Forecasting**
First Floor	0.952	0.882	42.6	0.011
Second Floor	0.954	0.887	52.9	0.015
Third Floor	0.955	0.885	42.9	0.013
Whole Building	0.954	0.885	138.4	0.039

3.7.2 Medium Office Summer Cooling Estimation

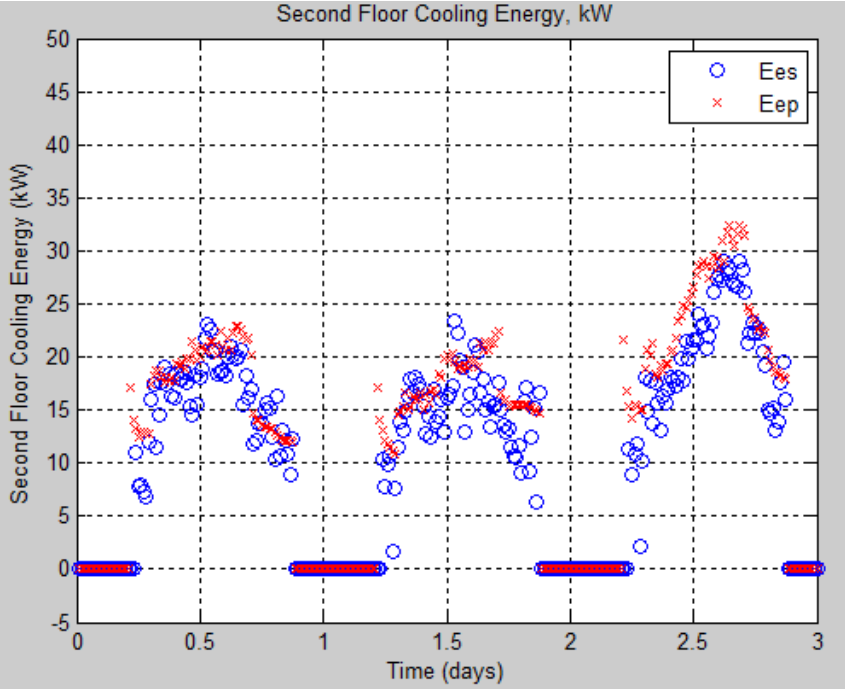
The forecasting accuracy and speed are summarized in Table 3-7, which illustrates that the forecasting results capture the trend of real energy consumption "measurement". The accuracy and speed are still acceptable for MPC, even though they are not as good as those for the small building, because medium building has more disturbances and it requires longer calculation time to include all these disturbances into the model. It is illustrated in Figure 3-13 that the on-line estimation model underestimates the cooling energy consumption during the HVAC starting up and shutting down periods. Comparing to the forecasting results of the small building, the forecasting results in this medium building has lagers errors. That is because there are more disturbances in the medium building than in small buildings. However, the on-line model still capture the overall trend of the real energy consumption and the accuracy is still above 88%. Next step of this research will focus on updating forecasting results based on real measurements when it is necessary to improve the accuracy, especially at the high dynamics period.



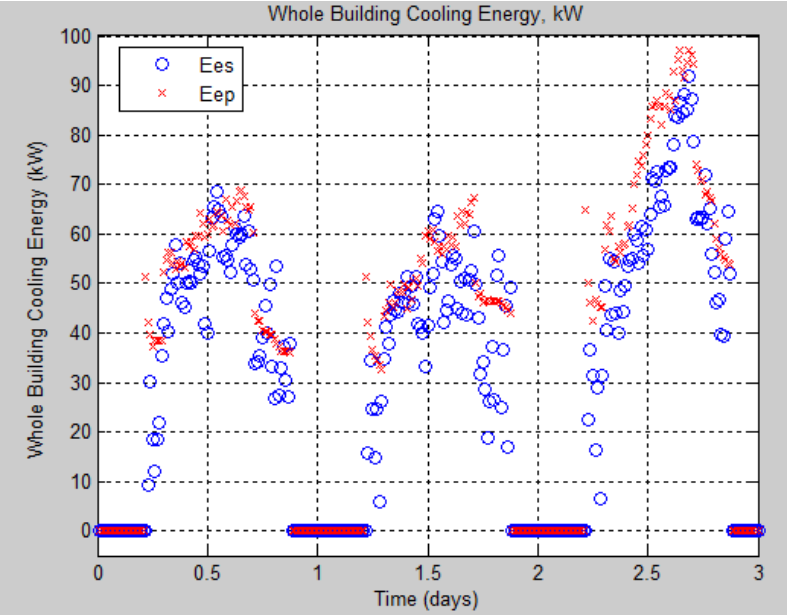
a) First floor model training results



b) First floor cooling energy forecasting



c) Second floor cooling energy forecasting



d) Whole building cooling energy

Figure 3-13. Medium building cooling energy estimation forecasting

3.8 Conclusion and Future Work

This study introduced a novel systematic methodology for on-line building energy estimation model development and validation. A system excitement scheme and a Matlab-BCVTB-EnergyPlus testbed was developed and validated for system identification model development. The excitement scheme is able to guarantee enough data at the high frequency and low frequency around the building operation range, which can be applied in any other system identification models for building energy simulation. Frequency response function model with realized by power spectral density model was implanted to forecasting building load and energy consumption. This on-line building energy model can achieve over 95% forecasting accuracy within one second in a small building case, which is suitable for any on-line building operation and MPC model. This modeling approach has also been applied in a medium office building. Limited to the calculation speed, the cooling energy forecasting accuracy is still around 88%.

This feasibility study demonstrates that using SI techniques can result in high fidelity online building energy forecasting model, which does not require extensive model development effort or training data over a long time period. However, only limited building types are examined here. A methodology to apply the SI techniques in different building types is needed (and discussed in Chapter 4). An experimental plan to validate the proposed methodology in a real building setting is needed (and discussed in Chapter 7).

4. CHAPTER 4 On-line Building Energy Forecasting: System Identification Methodology

4.1 Background and Motivation

As already discussed in Chapter 3, building energy forecasting model is critical to model based advanced building control and operation strategies. The accuracy and speed of a forecasting model are two most important evaluation criteria. As discussed in Chapter 1, existing energy forecasting models, no matter, white box (physics based) models, black box (data-driven) models or grey box (hybrid) models, all have their limitations: white box models have thousands of parameters to determine and the calculation time is too slow; pure data-driven black box models often require long training period and are bounded to building operation conditions; on the other hand, creating even a simplified grey box model using literature-reported approaches is still time consuming and needs expert knowledge. Therefore, this study proposes to develop a building energy forecasting methodology for building on-line building control and operation using system identification approach.

4.1.1 Building Energy System

Systems that strongly affect a building's energy consumption referred to as building energy systems, which includes building HVAC (heating, ventilation and air conditioning) system, lighting system and other internal equipment. The function of a HVAC system is to maintain the indoor thermal comfort (temperature and humidity) and indoor air quality. Due to the linearity of lighting and internal equipment, forecasting these systems energy consumption is somewhat straightforward. Therefore, this mainly focuses on the building HVAC system. A HVAC system's energy consumption is determined by the heating/cooling loads that the heating/cooling loads that the system is to satisfy and the system's efficiency as expressed in Eq. 4.1. A building's

heating/cooling loads are the heating/cooling energy that the HVAC system needs to provide to maintain a desired indoor thermal condition (temperature and humidity). Factors that affect a building's heating/cooling loads are graphically illustrated in Figure 4-1. The factors include solar heat gains (Q_1), internal heat gains (Q_2), ventilation air load (Q_3), and heat gains through envelope (Q_4). The building load can be calculated using Eq. 4.2, based on these factors.

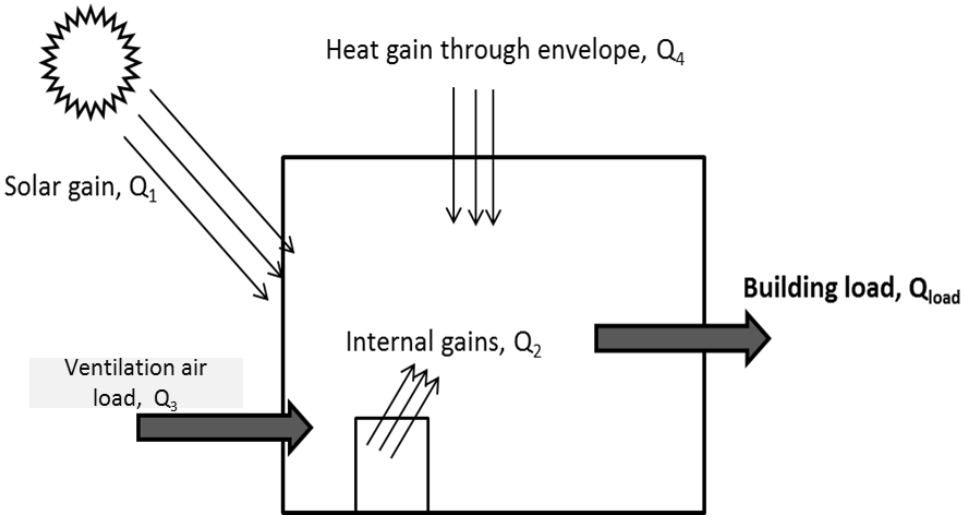


Figure 4-1. Factors for building heating/cooling loads

Moreover, since solar heat gain is determined by the solar irradiance (I_{sol}) and building envelope's material properties, such as transmissivity, absorbability, etc.; internal gains are determined by the internal equipment power (P_e) and its operation schedules (S_e). Ventilation air load is determined bu the ventilation flow rate (V_{oa}), outdoor air temperature (T_o) and humidity, and indoor air temperature (T_i) and humidity. Heat gain through envelope is determined by the outdoor air temperature (T_o), indoor air temperature (T_i) and overall envelope heat transfer coefficients (α_e). Therefore, the building load can be further calculated by Eq. 4.3.

$$E_{HVAC} = f(Q_{load}, C_{hvac}) \quad \text{Eq. 4.1}$$

$$Q_{load} = \sum_{i=1}^{n} Q_i \quad \text{Eq. 4.2}$$

$$Q_{load} = f(I_{sol}, T_o, T_i, P_e, V_{oa}, \alpha_e) \quad \text{Eq. 4.3}$$

Therefore, the main inputs for a building energy forecasting model are building operation settings, such as temperature setpoints, internal equipment operation schedules, ventilation air flow rate, etc., outside weather conditions, including outdoor temperature, humidity, and solar irradiance, etc., building envelope properties, such as transmissivity, absorbability, etc., and HVAC equipment efficiency. Among these inputs, building operation settings, outdoor weather conditions are typically measured. The goal of a system identification process is therefore finding a mathematic representation between the inputs and building energy consumption.

4.1.2 Building Energy On-line Forecasting Challenges

Although input variables that strongly affect a building's energy consumption can be determined easily based on the above physical analysis, the functions between these inputs and the parameters in the functions are complex and hard to obtain. Unlike other systems, such as aircraft systems, which have limited models and configurations, building systems have large variety including both building envelope variety, HVAC system variety, and internal condition variety. Therefore, it is extremely difficult to develop a universal energy forecasting model for all of the different building types through limited experiment studies. On the other hand, even developing an energy forecasting model for a specific building is not an easy task due to the complexity of the building energy system. As mentioned earlier, developing physics based models, even a grey box model, requires large number of parameter determination and customization, which result in expensive engineering efforts. The building energy systems usually are nonlinear, so their system order and parameters often vary with different weather and

internal conditions. Collecting existing building operation data for data driven models are time consuming and may not have enough data available. Therefore, this study aims at developing a systematic methodology that can be applied to different kind of buildings to forecasting their energy consumption for building's on-line control and operation through system identification approach.

4.1.3 System Identification Procedure

During a system identification process, model structure, model order, and model parameters are three most important features that need to be determined. They can be determined from the system operation data either from passive collection or active experiment. The overall procedure of building energy forecasting model development is illustrated in Figure 4-2.

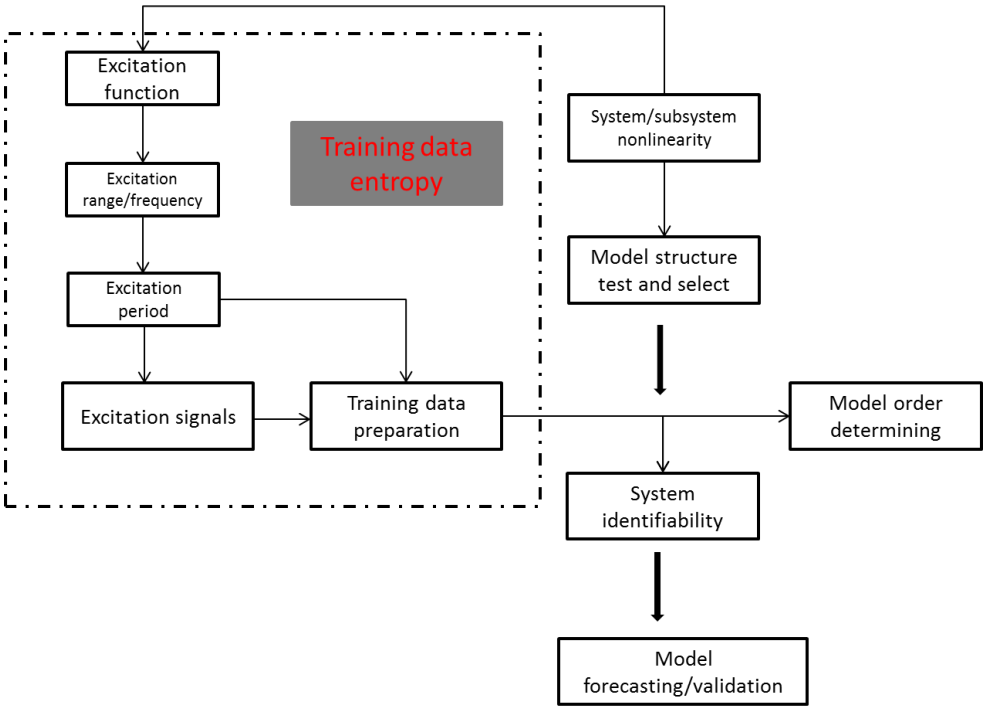


Figure 4-2. System identification based building energy forecasting model development

System identification model structure selection is the fundamental of system identification model development, which plays a significant role in model's accuracy and robustness. Usually, most of the model structures for parametric system identification are selected based on the prior knowledge or trial and error processes. However, this study proposes to develop a building energy system thermal characteristics test plan as guidance for model structure selection prior information. Besides model structure, model order is also critical for energy forecasting model. From previous studies [149, 174], there are different methods to determine the system order from system experiment data. After the model structure and order are determined, the model parameters then can be identified using parameter identification methods. The detailed methodology of system structure selection, order determination and parameter identification will be discussed in following sections.

In the end, this study will try to create a general building energy forecasting model mythology, including determining which building/system should use physics based model and/or data driven model, how to integrate models from different nature to form an effective system model, how to determine the model hierarchy for a complex building cluster (nonlinear with very different response time scale).

4.2 Building Energy System Prior Information Test: Nonlinearity Test

A system's nonlinearity affects the system identification structure selection. Therefore a system nonlinearity test will be designed and applied to obtain necessary information for system identification.

Generally, the nonlinearity of building energy system is dependent on, but not limited to, building size, floor are, HVAC system type, and HVAC system size. The physics theories for building energy system are very complicated and with a lot of disturbances. It is very difficult to

determine the system characteristics of building energy system just from theoretical analysis. Therefore, the building energy system nonlinearity test will be conduct on different virtual buildings with different sizes, floor areas and equipment types. The different cases of nonlinearity simulation test are tabulated in Table 4-1.

Table 4-1. Building energy nonlinearity simulation test cases

Building size	Building Floor	HVAC system
Small	1	Unitial or CAV
Medium	3	CAV or VAV
Large	5+	VAV

4.2.1 Nonlinearity Test Method

System nonlinearity is the most important characteristic, especially for nonparametric method, which determines linear or nonlinear model structure required to model the building energy system. As a result, the essential task at the beginning is to decide from system input and output measurements whether the system under study is linear or nonlinear. Usually, there are many methods for nonlinearity test, such as spectral density method, correlation method, Gaussian distribution test and so on [178]. These methods determine system's nonlinearity by applying different inputs and investigating the system outputs together with corresponding inputs.

As introduced previously, the nonlinearity of building energy system comes from two major parts: building envelope and building equipment. Therefore, three different tests are designed in this project. The first one is for the building envelope, the second one is for the building equipment, and the last one is for the whole building. All these three test are firstly intended to be conducted upon detailed simulation models. In the existing building energy simulation models, there are building envelope models for building load forecasting, building heating/cooling

equipment models for heating/cooling equipment operation simulation, and whole building model for overall building energy system simulation.

The overall testing approach is to apply different input signals into the system and examine its outputs, as shown in Figure 4-3. The general unchanging inputs are weather conditions, room occupancy, and other variables. They are either boundary conditions or predetermined variables. Even though they will affect the behavior of building energy system, they cannot be changed in the system testing period.

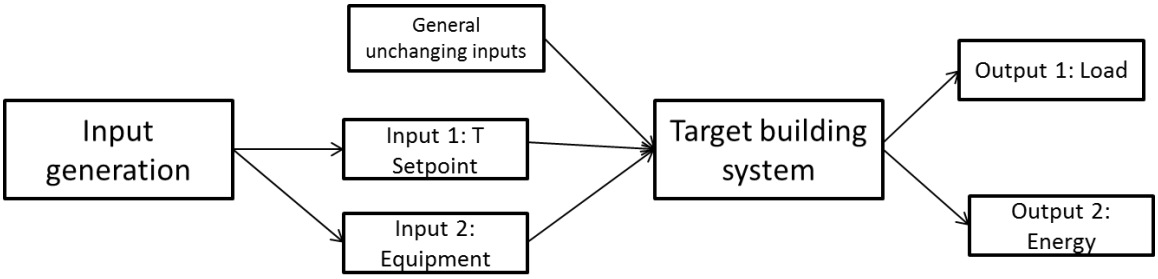


Figure 4-3. System nonlinearity test procedure

In the building envelope testing, the target system is the building envelope. The nonlinearity of building envelope is tested by applying different temperature setpoints and non-HVAC (heating, ventilation, and air conditioning) equipment schedules, such as lighting, computer. In the second step, the target system is the building heating/cooling equipment. The nonlinearity of building HVAC system is by applying same temperature setpoints and non-HVAC equipment schedules, while changing heating/cooling equipment settings and examining the building cooling or heating energy consumption. The overall nonlinearity of the building energy system is tested by applying different temperature setpoints and non-HVAC equipment schedules and directly examining the building cooling or heating energy consumption. The input and output of these three tests are summarized in Table 4-2.

Table 4-2. System identification model forecasting accuracy and speed for medium building

Target System	General input	Control input	Checking output
Building Envelope	Outdoor weather, occupancy…	Temperature setpoint, internal non-HVAC equipment schedule	Building heating/cooling load
HVAC System	Outdoor weather, occupancy…	Building heating/cooling load	HVAC equipment energy consumption
Building Energy system	Outdoor weather, occupancy…	Temperature setpoint, internal non-HVAC equipment schedule	HVAC equipment energy consumption

4.2.2 Nonlinearity Testing Input Signal Design

Buildings usually operate within a narrow temperature setpoint range, which is not enough to study the full scale system characteristics. Therefore, the inputs for the nonlinearity test are designed to operate the building in different conditions, which will generate the operation data in the full range of operation data for system characteristics analysis. For example, the sum of sin function designed for system excitation in section 3.3.4 can also be used in this section for system nonlinearity test. Other input signals, such as pseudorandom binary signal, Gaussian distribution signal will be applied, if the sum of sin function cannot generate necessary operation data.

4.2.3 Nonlinearity Evaluation

It is believed that a system's nonlinearity is one of the most important characteristics for a system's model development, especially for nonparametric methods [149]. In this study, a magnitude squared coherence based method for system nonlinearity test [149] is adopted. This method is based on the cross-spectral density of the inputs and outputs:

$$C_{xy} = \frac{|G_{xy}|^2}{G_{xx}G_{yy}} \qquad \text{Eq. 4.4}$$

Where, the magnitude squared coherence (C_{xy}) estimate the power transfer between input and output to estimate the causality between system input and output. G_{xy} is the cross power spectral density between system inputs (x), such as outdoor air temperature, and system output (y), such

as building energy consumption. G_{xx} and G_{yy} are the auto power spectral density of x and y, respectively. They can be estimated from the Fourier transformation of the auto-correlation of the inputs (R_{xx}) and outputs (R_{yy}), and the cross-correlation between inputs and outputs (R_{xy}). The equations for the transformation are presented as Eq. 4.5 to Eq. 4.7:

$$G_{xy} = \frac{1}{N}\sum_{\tau=1}^{N} R_{xy}(\tau)e^{-j\frac{2\pi k\tau}{l}} \quad \text{Eq. 4.5}$$

$$G_{xx} = \frac{1}{N}\sum_{\tau=1}^{N} R_{xx}(\tau)e^{-j\frac{2\pi k\tau}{l}} \quad \text{Eq. 4.6}$$

$$G_{yy} = \frac{1}{N}\sum_{\tau=1}^{N} R_{yy}(\tau)e^{-j\frac{2\pi k\tau}{l}} \quad \text{Eq. 4.7}$$

Where, l is sampling window length of the spectral density analysis. N is the number of sampling windows. The whole purpose of the Fourier transformation is to convert the signal into frequency domain to capture the system dynamics. R_{xy} is calculated in (Eq.5), R_{xx} and R_{yy} are calculated in Eq. 4.8 and Eq. 4.9.

$$R_{xy}(\tau) = \frac{1}{l}\sum_{i=1}^{l-1} u(i)\, y^T(i+\tau) \quad \text{Eq. 4.8}$$

$$R_{xx}(\tau) = \frac{1}{l-\tau}\sum_{\tau=1}^{l-\tau} x(j)x^T(j+\tau) \quad \text{Eq. 4.9}$$

During an auto-correlation process, the similarity between an observation and the same variable with a time lag is analyzed to discover similar patterns in a signal. Cross-correlation is a measure of similarity between one signal and another signal with lag time. For example, cross-correlation can be used to examine the similarity between temperature setpoint and cooling energy consumption. The power spectral density describes how the power of a signal is distributed over different frequencies. The cross power spectral density can be calculated from the

Fourier transformation of the cross-correlation between two signals, and the auto power spectral density can be calculated from Fourier transformation of the auto-correlation of input signals.

In this nonlinearity evaluation process, the entire analysis period (e.g. one day) will be divided into multiple moving Welch's overlapped l -long (e.g. 6 hours) segments. The nonlinearity evaluation process (Eq.2 – Eq.4) will be conducted in each segment. The overlapping portion in this study is chosen as 50%. That means the 6-hour sampling window will move forward for 3 hours at each time. The sampling window and excitation injection are illustrated in Figure 4-4.

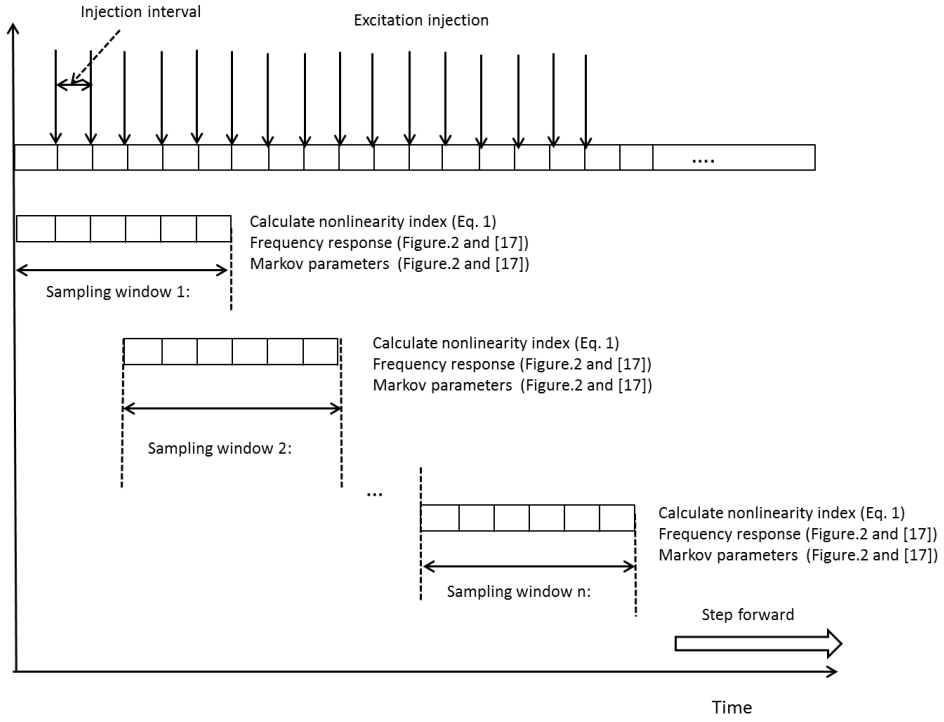


Figure 4-4. Data sampling window and excitation injection during a nonlinearity evaluation and system identification process.

The reason why overlapping segments window are used here is that most "window" functions afford more influence to the data at the center of the set than to data at the edges, which represents a loss of information. To mitigate that loss, the individual data sets are commonly

overlapped in time. During the nonlinearity evaluation, the C_{xy} will be evaluated at each time step (measurement) for each sampling window. In the initial trying case, the time step is 15 minutes. So there will be $\frac{60}{15} \times 6 = 24$ nonlinearity indexes for each sampling window. As the sampling window moves forward, the total number of sampling windows will be $\frac{24}{3} \times 7 = 56$. Where 24 is the hours per day, 3 is the sampling window moving speed, and 7 is the number of days in the initial trying cases. Therefore, the final nonlinearity indexes will be:

$$\overline{C_{xy}}(i) = \frac{\sum_{n=1}^{n=N} C_{xy,n}(i)}{N}$$

Where, $\overline{C_{xy}}(i)$ is the overall nonlinearity index at the time step I, $C_{xy,n}(i)$ is the nonlinearity index for the n^{th} sampling window at the time step i, and N is total number of sampling window.

If $C_{xy} = 1$, then the system is a linear system. Suppose $y(t) = h(t)x(t)$ is a linear system, the nonlinearity index can be calculated in Eq. 4.10.

$$C_{xy} = \frac{|G_{xy}|^2}{G_{xx}G_{yy}} = \frac{|H(f)G_{xx}(f)|^2}{G_{xx}^2(f)|H(f)|^2} = \frac{|G_{xx}(f)|^2}{G_{xx}^2(f)} = 1 \qquad \text{Eq. 4.10}$$

If $0 < C_{xy} < 1$, then the system a is nonlinear system. The prove of this statement can be found in [149]. And the closer the C_{xy} is to 1, the more the system behaves like a linear system. Details about the system nonlinearity test, including testing signal generation and testing results are discussed in section 4.5. The sampling window length is initially chosen as 6 hours, but is later updated based on the system characteristics.

As described above, this building energy system nonlinearity study will be conducted as following steps:

a) Categorize buildings into light, medium, and heavy buildings. The detailed categorizing factors are building story, floor area, and thermal response time. Building thermal response

time can be estimated by changing room operation settings (such as temperature setpoint) and measuring the time for the building to reach to new setting from the previous conditions. The estimation model has been discussed in Eq. 3.6. Usually, heavy buildings have longer response time than medium and light buildings. Therefore, two response time indexes will be identify to distinguish light, medium, and heavy buildings based on the simulation results.

b) Nonlinearity test for different buildings in each category. Similar to response time, different buildings have different nonlinearity. Therefore, the typical nonlinearity index for each category will be identified.

c) Divide the whole building energy system for building envelop and building HVAC system, and conduct the nonlinearity test for building envelop and building HVAC system individually.

4.3 Building Energy System Prior Information Test: System Response Time Test

Besides system nonlinearity, system thermal response time is another critical factor in determining system identification methodology, especially when determining the excitation plan and sampling window. The excitation signal generation frequency should be calculated based on the system response time (Appendix A), and the signal injection interview and sampling window should be larger than the response time to allow the system become stable.

System thermal response time is a measure of how quickly the system responds to an input change. System response time is usually measured by experiments. For example, if an response of a dynamic system can be expressed as [149]:

$$x(t) = (\alpha/T)e^{-t/T} \quad \text{Eq. 4.11}$$

Where, T is the response time constant, α is a state parameter, also known as measurement in the test. The response time of a system measurement, $x(t)$, to reach 95% of its final steady state

value after the input change, is defined as $T_{0.95}$. In this study, the building zone temperature is chosen as the measurement in this response time experiment. Since both the HVAC system capacity and the building thermal mass could affect its zone temperature response time, two tests are performed to evaluate the response time respectively. The first one is to change the zone temperature setpoint after the building zone temperature has reached a steady state, and then measure the time between the beginning of the setpoint change and when the zone temperature reaches the 95% of the new setpoint. This response time is a combination of both building thermal mass and HVAC system capacity. The other one is to switch off the HVAC system at night, when weather disturbances are minimal, and measure the time that the zone temperature takes to decrease to a steady state (or nearly a steady state). The nearly steady state is defined as less than 0.5% of the state change in 15 minutes. This second test evaluates the building thermal mass when the weather disturbances are minimized. The detailed testing procedure and results of the system nonlinearity and response time tests will be presented in section 4.5.

4.4 Building Energy System Identification Model Development

In this section, the detailed methods for system structure selection, order determination, and parameter identification are introduced.

4.4.1 Building Energy System Identification Model Structure Selection

Generally, model structure selection for data-aided system identification is an underdeveloped field [149]. There are two category model structures: time domain model and frequency domain model. The frequency domain model refers to the mathematical functions for the system with respect to frequency, rather than time. The most benefit of frequency domain model is to lower down the model order, even convert the nonlinear model to linear model, which is much easier to solve. However, there has not been much discussion about which model

structure is better for a building, nor frequency domain methods used in building energy forecasting. Therefore, this study will try to explore the use of frequency domain method in building energy forecasting and provide system structure selection recommendations for model energy forecasting model development.

To better recommend the model structure for building energy forecasting, this project also plans to compare the performance of different potential models, such as Frequency response model developed in Chapter 3, RC model, and different pure data driven models.

4.4.2 Building Energy System Identification Model Order Determination

Unlike model structure selection, there are many different ways to estimate a system's order based on data analysis:

1. Testing rank of covariance matrix

   Suppose the order of a system is "n", and let

$$\varphi_s(t) = [-y(t-1) \dots . -y(t-s) \;\; u(t-1) \dots u(t-s)]^T$$

   Where, $\varphi_s(t)$ is a testing vector containing both outputs, y, and inputs u. Then the covariance matrix will be:

$$R^s(N) = \frac{1}{N}\sum_{t=1}^{N} \varphi_s(t)\, \varphi_s^T(t)$$

Therefore, R will be nonsingular for $s \leq n$ and singular for $s > n$ [149]. Hence system order can be calculated by determining s.

2. Testing correlating variables

The order-determination is to include the right number of variables in the model. This problem can be solved by checking whether a new variable has any contribution for the output. This contribution can be measured by the correlation test between system outputs and inputs.

3. Testing rank of information matrix

Information matrix is a measure of the amount of information that a signal (x) contains about an unknown parameter (θ):

$$I_n(\theta)_{i,j} = -E\left[\frac{\partial^2 \ln f(X|\theta)}{\partial\theta_i \partial\theta_j}\right]$$

The rank of the information matrix, I_n, is the order of the system.

All these three methods are planned to be used and compared in the future system order determining study under different situation. An overall system order determination strategy will be recommended based on the comparison studies.

This information matrix is also related to the system identifiability. System identifiability is whether the identification procedure will yield a unique value of the parameter (θ) and whether the final model is representing the true system. A model structure $F(z, \theta)$ is locally structurally identifiable at $\theta^* \in \Theta$, if for all θ_1 and θ_2 in the neighborhood of θ^*, then for all z [179]:

$$F(z, \theta_1) = F(z, \theta_2) \Rightarrow \theta_1 = \theta_2$$

It is important to note that even a model is structurally identifiable but it might not be data dependent identifiable if the input data is low quality:

$$F(z, \theta_1)u_k = F(z, \theta_2)u_k \Rightarrow \theta_1 = \theta_2$$

Where, u_k is system input and model structure $F(z, \theta)$ is data dependent identifiable [179]. Therefore, in order to generate high quality system training data, a system excitation plan is developedd. To test the data-dependent system identifiability, Fisher matrix, *M*, bounds the covariance of the parameter estimation error according the Cramer-Rao inequality [149]. Suppose $\varepsilon_k(\theta) = y_k - \hat{y}_k$ and $\varepsilon_k \sim N(0, \Sigma_e)$, then *M* is defined as:

$$M = \sum_{k=1}^{N} \left[\frac{\partial \varepsilon_k(\theta)}{\partial \theta} \Sigma_e^{-1} \frac{\partial \varepsilon_k(\theta)}{\partial \theta}^T \right] = \sum_{k=1}^{N} \frac{\partial \hat{y}_k(\theta)}{\partial \theta} \Sigma_e^{-1} \frac{\partial \hat{y}_k(\theta)}{\partial \theta}^T \quad \text{Eq. 4.12}$$

However, the Fisher matrix cannot be used to select the system model structure, because it need structure model as prior information for calculation. The Fisher matrix usually uses to identify the model parameters, and guide to design experiment for training data collecting.

4.4.3 Building Energy System Identification Model Training Data Generation

Most of the building energy systems are typically only operating within a very small range of settings, for example, temperature setpoints, internal equipment schedules. Usually, the passive model training methods are using regular building operation data for certain days. The only way to improve the training data is to increase the training period time. Considering that an active training method is developedd in this study, which not only controls the length of training period but also improves the quality of training data. The active training method is realize by optimizing the excitation signal based on data information entropy and optimal experiment design theories. The excitation signals are needed to satisfy key theoretical assumptions on reliable statistical identification – persistent exciting signals [180]. Three basic "plant-friendly" excitation signal constraints were presented by Braun et al. [159]:

1. Keeping minimum deviations in control signal;
2. Implementing a signal of short duration to minimize the amount of off-spec product;
3. Keeping move sizes small to satisfy actuator constraints and minimize "wear and tear" on process equipment.

Pseudo-random binary signals, sum of sinusoids and multilevel pseudo-random signal are the three most common excitation signals in system identification. For dynamic systems, system excitation experiment design includes choice of inputs and outputs, test signals, and sampling

intervals. A preliminary study on excitation signal for training data generation has been introduced in section 3.3, where sum of sinusoids is used. Crest factor (Eq. 3.7) is applied to determine the frequency and phase of the signal. The system identification model developed on the training data achieved over 90% accuracy. Therefore, this sum of sinusoids function will also be used in future studies, while, other functions will also be tested if it is necessary.

The experiment design for system identification includes determining input signals, measurements, sampling intervals and how to manipulate measurements. From informative excitation theory, the excitation signals should be persistence exciting, and the order of excitation function should be equal to the number of parameters to be estimated in the system.

The sampling interval is also critical to system identification model performance and identifiability. Fast sampling leads to high-frequency bands problem and increase calculation burden. On the other hand, if the sampling interval is longer than the system natural time constant, the sampling data's variance increases drastically. The suitable sampling interval lies in the range of tem times the bandwidth of the system [149].

4.4.4 Building Description

To evaluate how different building size, envelope, and HVAC systems affect the SID process, two building sizes with different HVAC systems are used in this study as the objects. One is a small-size single story office building (same as that used in [181]). It has five conditioned zones and an unconditioned attic. The total floor area is 510 m^2. The window-to-wall ratio is approximately 21.2%. The overall U-factor of its single pane windows is 3.4 W/m^2K with a solar heat gain factor of 0.36. The overall U-factor of the external envelopes is 0.68 W/m^2K. The R-value and solar absorptivity value of the roof are 5.1 W/m^2K and 0.9, respectively. The HVAC

systems used in this building are constant-air-volume (CAV) air handling units (AHUs) with direct expansion (DX) coils.

The other building is a three-story office building, and each floor has five conditioned zones. The total floor area is 4982 m^2. The window-to-wall ratio is around 33%, and the U-factor of these windows is 3.3 W/m^2K with a solar heat gain factor of 0.36. The R-value and solar absorptivity value of the roof are 0.33 W/m^2K and 0.7, respectively. This building uses packaged multi-zone variable-air-volume (VAV) with electric reheat.

Very detailed physics based building simulation models using EnergyPlus [177] are used in this study in lieu of real building systems. The EnergyPlus models are developed by U.S. DOE and are validated by real field data [177]. In the small building model, the coefficients of performance (COP) of the DX coils is modeled using quadratic equation [177]. In the medium building model, the performance of the VAV system is modeled using second-order and third-order polynomial equation [177]. The location of these two building is selected as in Philadelphia, PA, USA for this study. Typical Meteorological Year (TMY) weather data file provided by the DOE is used as the weather input. A virtual building system identification emulator [182] is used here to simulate the building operation and apply the introduced system identification schemes.

4.4.5 Training and Testing Condition

The two buildings, as described in section 4.4.4, are simulated for three time periods under Philadelphia TMY data: 1) from August 1st to August 7th ; 2) from August 1st to August 14th ; 3) from August 22nd to August 28th; and 4) from July 1st to September 30th. Simulated building operation data during time period 1 is used as training data for the SID model. Simulated data during time period 2, which is longer than time period 1, is used as the training data for other models in the comparison study (described in Section 4.7). Simulated data during time period 3 is

used for model adaptation (described in section 4.6) and model uncertainty comparison study (described in section 4.7.2.3). Simulated data during time period 4 is used as the testing data for all models.

Two sets of simulated operating data are generated in this study for the model development and comparison. One set of data assumes no measurement noise and is labeled as noise-free data in the following sections. The other set of data includes measurement noise by adding Gaussian distributed random white noise to each measurement. More details about the noise generation process will be introduced in section 4.7.2.

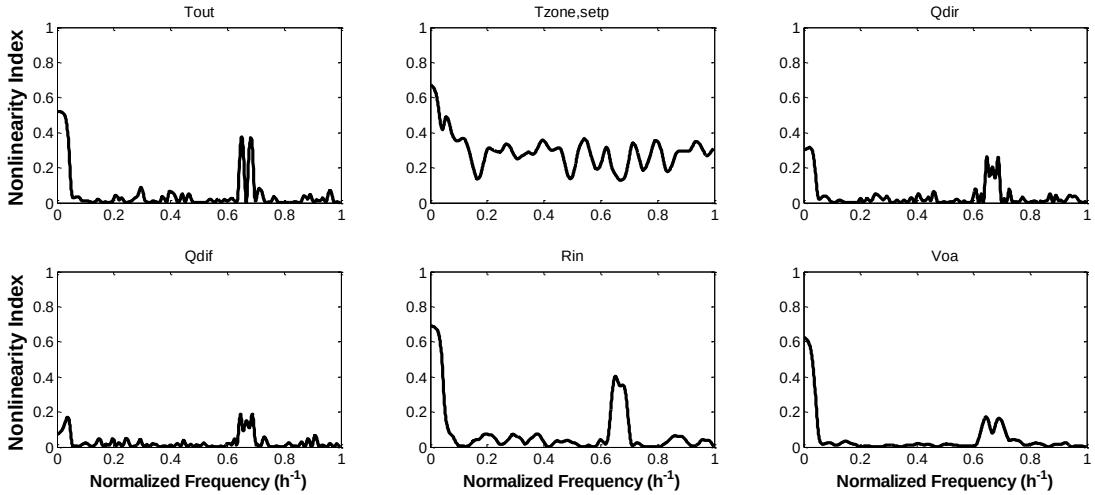
4.5 Building Energy System Characteristic Testing Results for System Identification

The results of the system nonlinearity and response time analysis are discussed in this section. These results are then used to adapt the SID model to improve the overall model performance.

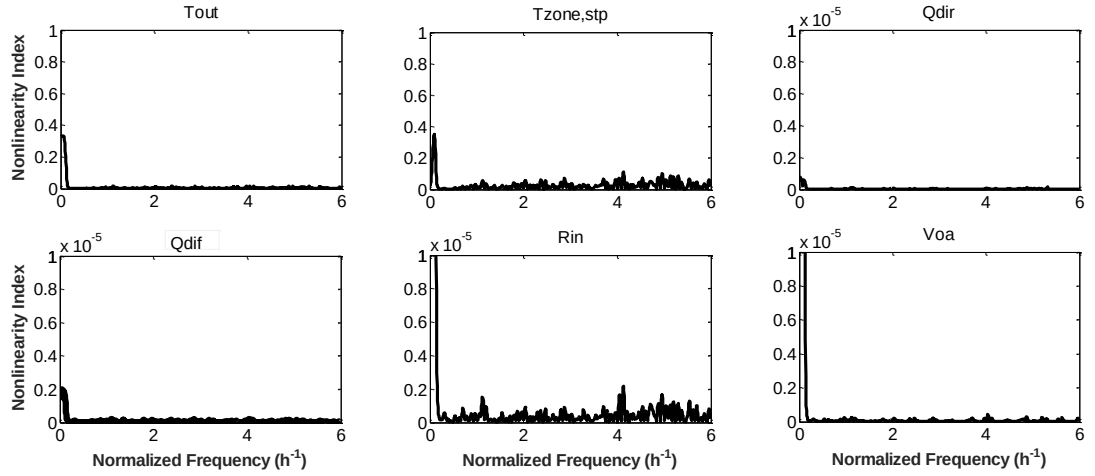
4.5.1 System nonlinearity test results

Using the system nonlinearity test methods introduced in section 4.2.1, the system nonlinearity between each system input and output are calculated for both of the two buildings. The testing signals are generated following the system excitation process described in section 4.2.2. Figure 4-5 summarizes the nonlinearity test results, where each subplot illustrates the nonlinearity between the output and a specific input, as a function of the input frequency. The nonlinearity index is calculated within the data sampling window (6 hours). As this figure shows, the nonlinearity between outdoor air temperature and the system cooling energy of the small building (Figure 4-5a) is around or above 0.2. This nonlinearity is around 0.5 for zone temperature setpoint. However, both of them are less than 0.001 in the medium building case (Figure 4-5b). Similar trends are observed for lighting/equipment schedule and ventilation rate. This indicates that the medium building is much more nonlinear than the small building. In order

to use the frequency Therefore, better excitation and data sampling schemes are needed to improve the performance of SID model in the medium building.



a. Small building nonlinearity test results



b. Medium building nonlinearity test results

Figure 4-5. Nonlinearity test results

4.5.2 Solar radiation delay factor

In addition, the nonlinearity indexes between the solar radiation (direct and diffuse solar radiation) and the system output are less than 0.001 in both small and medium building cases. It is well-understood that the impact of solar radiation on a building's energy consumption is often

delayed due to a building's thermal mass storage [183].Therefore, in order to capture this "delay" effect of the solar radiation, a variable correlation study has been conducted. In this test, the direct and diffuse solar radiation variables are delayed every 30 minutes until 3 hours. The correlations between the solar radiation and the cooling energy consumption at all the cases are summarized in Table 4-3 for small building and in Table 4-4 for medium building.

Table 4-3. Small building solar radiation delay correlation test results

Hour delay	**Whole building cooling energy**	
	Direct solar	Diffuse solar
0	0.482	0.517
0.5	0.507	0.541
1	0.513	0.555
1.5	**0.518**	**0.565**
2	0.496	0.564
2.5	0.478	0.558
3	0.449	0.586

Table 4-4. Medium building solar radiation delay correlation test results

Hour	**First floor cooling energy**		**Second floor cooling energy**		**Third floor cooling energy**	
delay	Direct solar	Diffuse solar	Direct solar	Diffuse solar	Direct solar	Diffuse solar
0	0.674	0.632	0.639	0.6	0.603	0.554
0.5	0.692	0.646	0.659	0.618	0.631	0.581
1	**0.702**	**0.652**	0.672	0.628	0.655	0.603
1.5	0.701	0.649	**0.677**	**0.630**	0.670	0.616
2	0.692	0.637	0.673	0.623	**0.677**	**0.620**
2.5	0.671	0.617	0.656	0.604	0.672	0.613
3	0.642	0.586	0.630	0.576	0.657	0.595

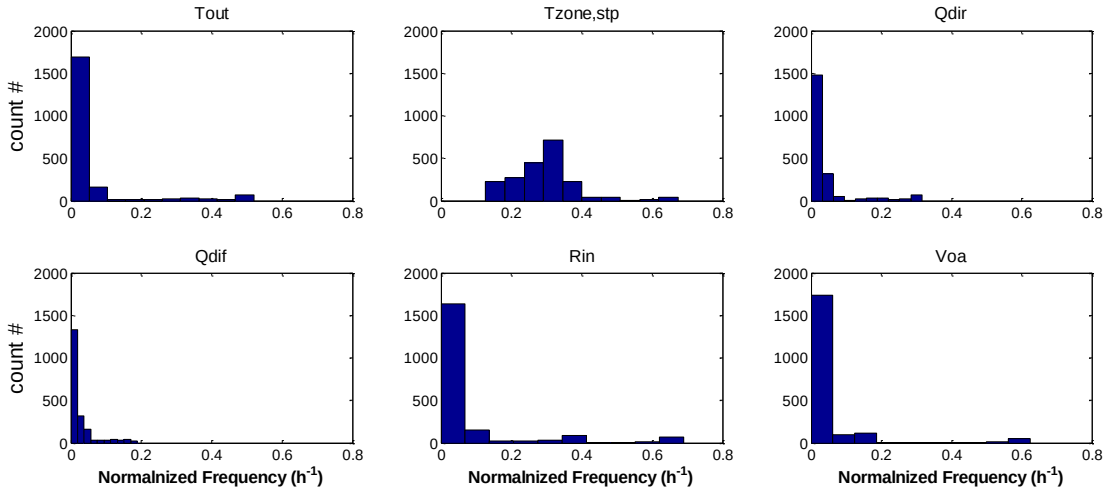
As Table 4-3 makes clear, the solar radiation variables at 1.5 hours delay have the highest correlation factor to the cooling energy consumption. It is shown in Table 4-4 that the solar radiation "delay" effect is different from floor to floor in the medium building, solar radiation variables at 1 hour delay have the highest correlation factor to the first floor cooling energy consumption, 1.5 hours delay to the second floor cooling energy consumption, and 2 hours delay

to the third floor cooling energy consumption. Therefore the solar radiation variables will be delayed according to these delay factors in the following model adaptation process.

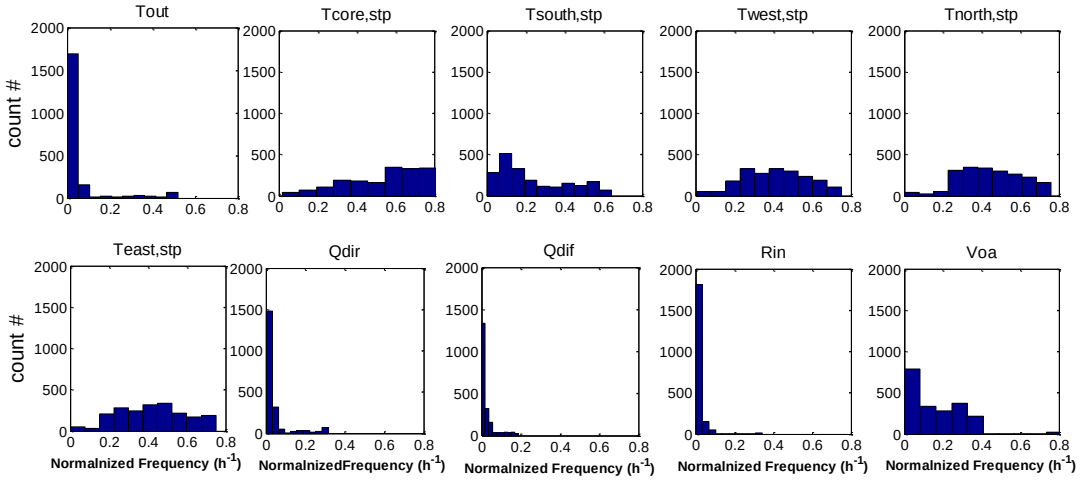
4.5.3 New excitation scheme and results

Considering the relationship between system inputs and outputs are different from system to system in one building, in order to excite the building more suitable to each system, different excitation signals are generated for each zone in the medium building, instead of using the master excitation signal for the whole building as that in the small building.

Based on the analysis above, another system nonlinearity test is then conducted using a new system excitation plan which considers the solar radiation time delay factors. In this new plan, the medium building is excited zone by zone with excitation signals at lower frequency range. The frequency distribution of the excitation signals (Tzone,stp, Rin) and other system inputs signals are plotted in Figure 4-6 for both small building and medium building case. In the small building case (Figure 4-6a), all the input signals except Tzone,stp are mostly distributed in the lower frequency range between 0 to 0.1, and the Tzone,stp are distributed between 0.1 to 0.3. Similar situation has been found in the medium case (Figure 4-6b). Therefore, the testing signals in the new excitation and model training period are mostly distributed in the lower frequency range, which matches the findings in the system nonlinearity test. Based on the new system excitation plan and system inputs, the system nonlinearity for both small and medium building are tested again.



a. Small building input signal frequency

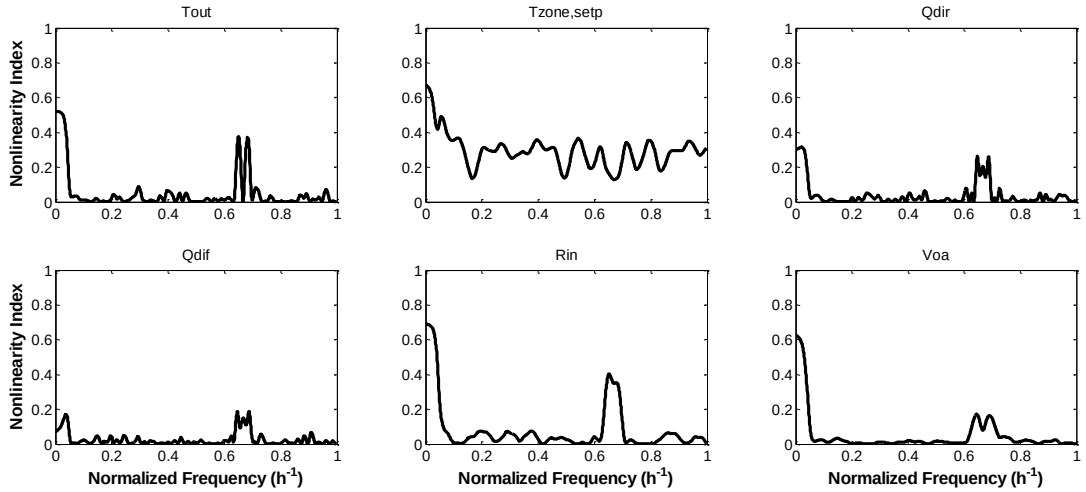


b. Medium building input signal frequency

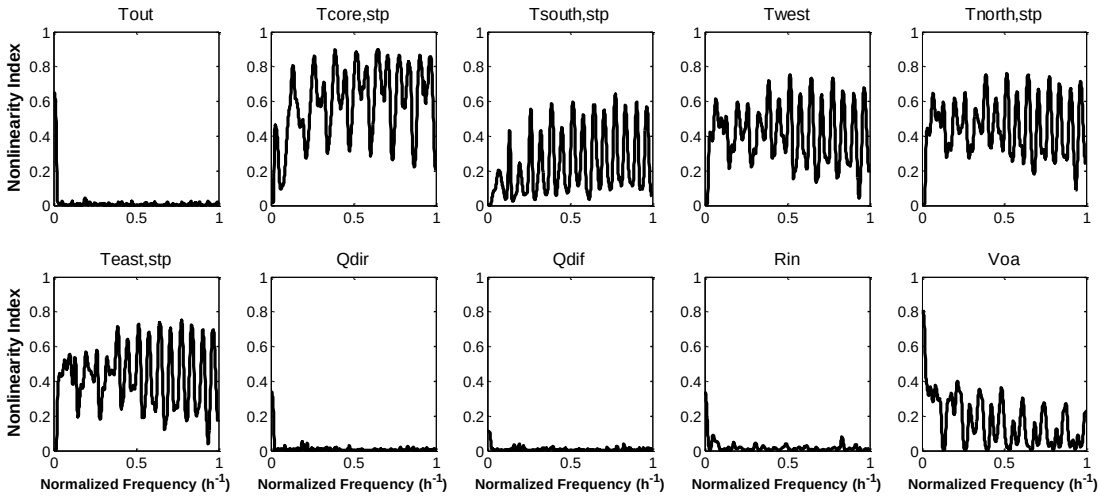
Figure 4-6. Building input signal frequency histogram

The results of this new test are shown in Figure 4-7. In this test, same nonlinearity evaluation approach is used for these two buildings with the updated excitation signals. The nonlinearity indexes for solar radiation (4^{th} and 5^{th} inputs) in the small building case (Figure 4-7a) is improved significant, comparing to those in Figure 4-5. The improvement of the solar radiation nonlinearity indexes in the medium building test is not as significant as that in the small building case. This means that the effect of solar radiation on the energy consumption in the medium building case is

more nonlinear. For the medium building (Figure 4-10b) the indexes for the zone temperature setpoints (2nd to 6th inputs) under the "zone-by-zone" excitation scheme are all between 0.5 and 0.8 for the four perimeter zones and between 0.2 and 0.8 for the core zone. This means that the system under this operation strategy is closer to a linear system comparing to the situation under the previous operation strategy. Following the results from this test, this new excitation scheme and solar radiation time delay are applied in the adapted SID model development.



a. Small building nonlinearity test results



b. Medium building nonlinearity test results

Figure 4-7. Building nonlinearity test results under new excitation

4.5.4 System response time test results

The system response time is tested using EnergyPlus model following the testing procedure introduced in section 4.3. As introduced, the system response time is tested in two different ways with HVAC system being on or off. Depending on the zone temperature range during the model training and forecasting period, three different cases are tested for the HVAC-system-on test (test 1). The three cases are summarized in Table 4-5.

Table 4-5. System response time test results

Setpoint change	**Temperature change time, $T_{0.95}$ (test 1)**		
	18 ºC - 22 ºC (22 ºC -18 ºC)	22 ºC -28ºC (28ºC -22 ºC)	18ºC -28 ºC (28 ºC -18 ºC)
Small	10.0 (9.5) min	8.0 (7.5) min	18.5 (9.0) min
Medium	33.5 (49.0) min	69.0 (31.0) min	88.0 (67.5) min
Initial temperature	**Temperature decay time, $T_{0.95}$ (test 2)**		
	18ºC	23 ºC	28 ºC
Small	47 min	39 min	23 min
Medium	67 min	48 min	36 min

For the HVAC-system-off test (test 2), which is conducted at night to take out the solar radiation disturbance, the building is firstly maintained at a designed temperature and then the HVAC system will be turned off. Three cases with different initial zone temperatures are designed in this test. The time that the building temperature takes to reach to stabilize after the system change is measured and also summarized in Table 4-5. System response time test results The results of the first test show that the system response time of small building is less than 10 minutes for most of the temperature setpoint changing cases in the small building, while those in the medium building case are over 45 minutes for most of the cases. The test results are reasonable, because the medium building has heavier envelope thermal mass, and the thermal capacity of the medium building is larger. Therefore it need more time for the HVAC system to

change the building temperature to the new setpoint in the medium building. This also explain the results of the second test that it also takes longer time for the zone temperature of the medium building to reach to the stable statues when the HVAC system is turned off. Therefore, based on the system response time test results, the system excitation frequency, excitation injection interval, and sampling window length should be adapted accordingly. The details about the model adaptation based on the system nonlinearity and response time test is discussed in section 4.6.3.

4.6 Building Energy On-line Forecasting Model Adaptation Results

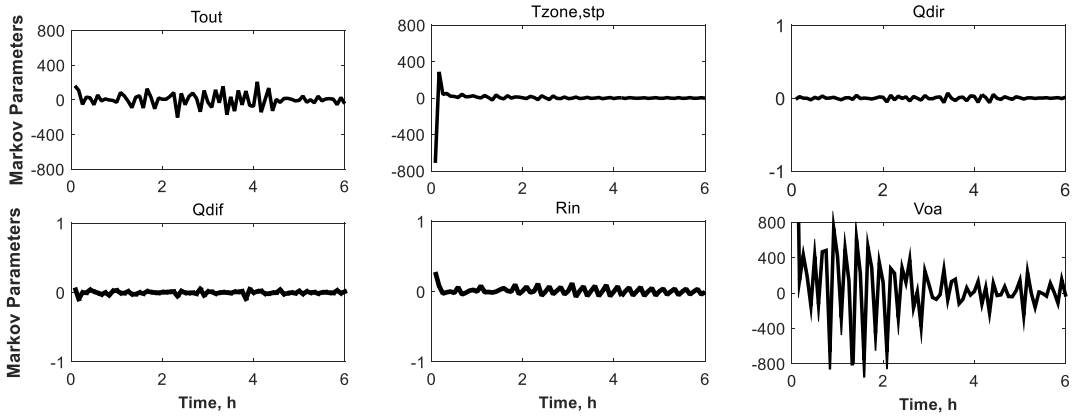
In this section, the SID model reported in section 3 is adapted for the medium building case based on its system characteristics. Firstly, the original SID (referred to as "pre-adaptation" hereafter) model is applied for both of the two buildings with varying HVAC systems. In this section, the performance of the pre-adaptation SID model, when used for the small and medium buildings, is summarized. Then how to adapt the SID model based on the system characteristics identified from Section 4.6 is discussed. The performance of the adapted SID models for both buildings is summarized at last.

4.6.1 Pre-adaptation

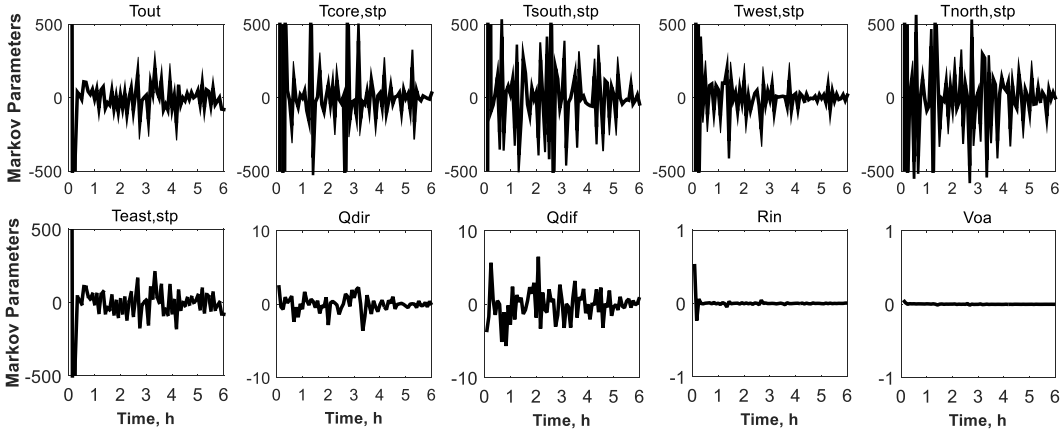
4.6.1.1 Markov parameters

As introduced in the SID model development section, the energy forecasting model is saved as a set of Markov parameters for each input, which are calculated in the training process at each 15-mintue time step within each 6-hour sampling window The Markov parameters of the SID models for these two buildings are plotted in Figure 4-8. In Figure 4-8a, the Markov parameters for "solar radiation" and "equipment schedule" (3^{th} to 5^{th} inputs) do not converge to zero at the end of the sampling period (6 hours) in the small building model. This indicates that the 6-hour sampling window is not long enough for this small building. The situation is different for the

medium building model, where all of the Markov parameters converge to zero at the end of the sampling window. Therefore, considering the truncation error, the sampling window lengths of these two buildings are adapted accordingly. The adaption study is introduced in section 4.6.3.



a. Small building Markov parameters



b. Medium building Markov parameters

Figure 4-8. Building SID model Markov parameters

4.6.1.2 Energy forecasting results before model adaptation

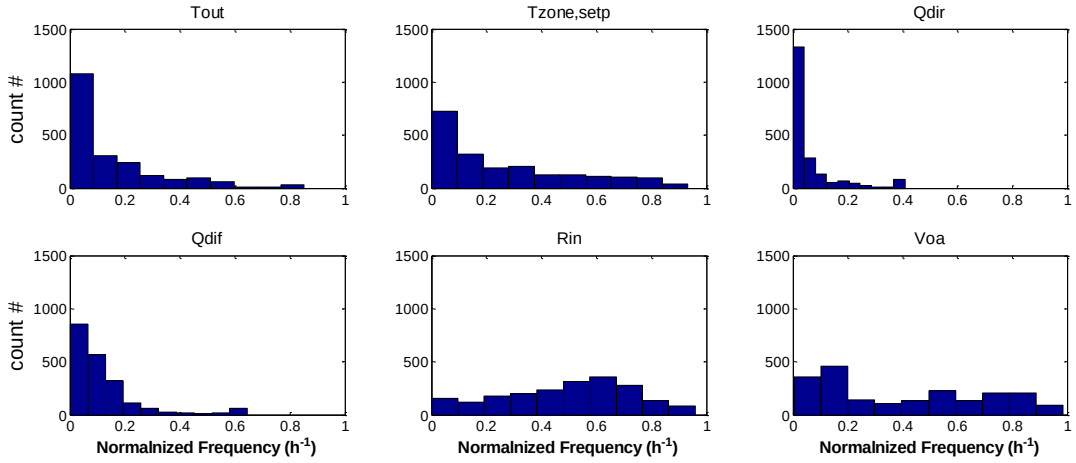
The energy forecasting results from the original SID model are discussed in this section. In order to evaluate the forecasting accuracy, three indexes are employed, namely, Coefficient of Determination (R^2), Root Mean Square Error (RMSE) and Normalized Root Mean Square Error (NRMSE):

$$R^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(\hat{x}_i - \bar{\hat{x}})}{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(\hat{x}_i - \bar{\hat{x}})} \quad \text{Eq. 4.13}$$

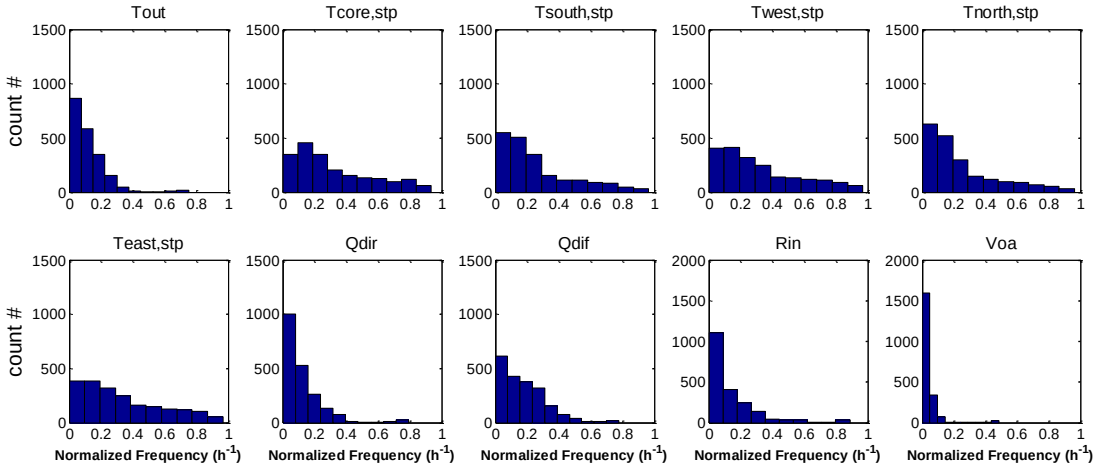
$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{n}} \quad \text{Eq. 4.14}$$

$$NRMSE = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{n}} \Big/ (x_{max} - x_{min}) \quad \text{Eq. 4.15}$$

Where x_i and $\hat{x}_i$ is the true and forecasting value; $\bar{x}$ and $\bar{\hat{x}}$ are the average of true and forecasting value, respectively.



a. Small building input signal frequency



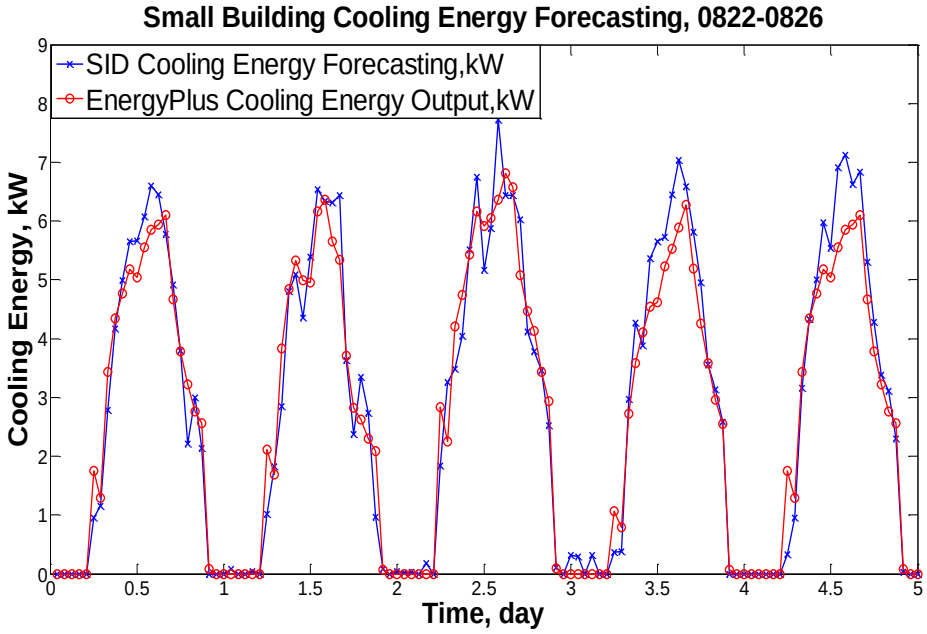
b. Medium building input signal frequency

Figure 4-9. Building input signal frequency histogram in forecasting period

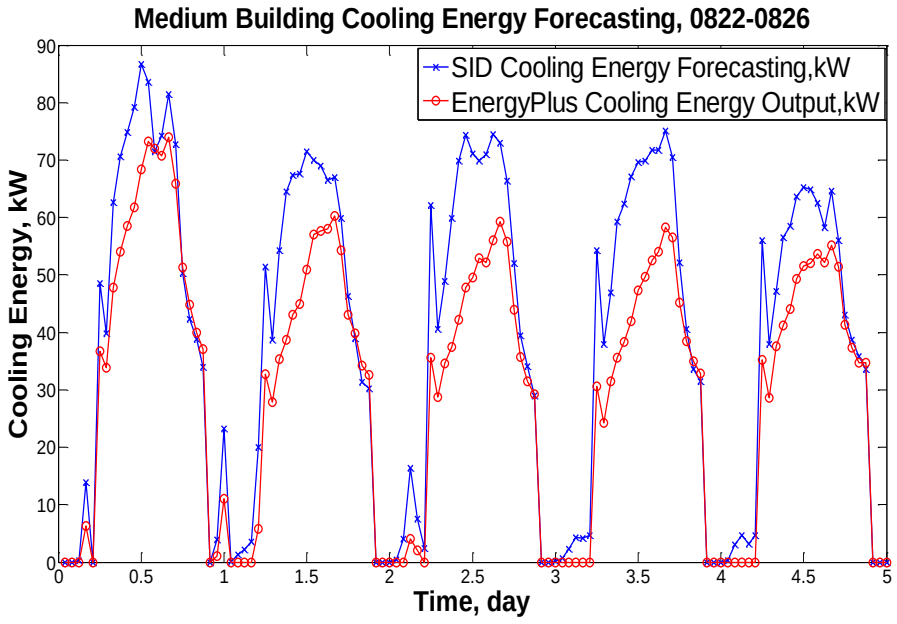
In section 4.5, the frequency distribution of system inputs are examined for the training data. In this section, these frequency distributions are re-examined to see whether the relationship between system output and inputs are also close to linear during a normal system operation. The input signal distributions during the forecasting period are plotted in Figure 4-9. In the plot for small building (Figure 4-9a), the input signals are mostly distributed in the frequency range between 0 and 0.2, except equipment schedule (R_{in}), and ventilation rate (V_{oa}). R_{in} distributes more in the frequency range between 0.6 and 0.8 and V_{oa} has higher distribution density in the following two ranges: 0 to 0.2 and 0.6 to 0.8. When comparing these ranges to their corresponding nonlinearity index distribution (Figure 4-7), it is found that all of these distribution ranges will yield a system nonlinearity index that is also closer to 1. Therefore, the normal input signals in the real building are more distributed in the range where the system behaves closer to a linear system. In the medium building case (Figure 4-9b), the signals are distributed more in the frequency range of 0 to 0.4, where the system nonlinearity index is also higher. Therefore, the SID model developed under the new excitation scheme matches well to the real building operation case, which will guarantee the model performance.

Again, as described in section 4.4.5, training data (August 1st to 7th) are used firstly to developed the pre-adaption SID model. The model is then tested in time period 3 (August 22nd to August 28th). During a model testing, the model is given an initial condition and then used to forecast building energy consumption (model output) for the entire testing time period. The detailed testing results for pre-adaption SID model are illustrated in Figure 4-10. Model forecasting accuracy and speed are summarized in Table 4-6. It is showed that, for the small building case, the pre-adaption SID model is able to achieve acceptable forecasting accuracy (R^2>0.95) within less than 1 minute calculation time. However, the performance of the pre-

adaption SID model in the medium building case is much worse (R^2<0.8). As a result, adaptation of the SID model to improve its performance is desired for the medium building case.



a. Small building energy forecasting



b. Medium building energy forecasting

Figure 4-10. Building cooling energy forecasting results before adaptation

Table 4-6. Unmodified SID model performance

Building	Calculation time	R^2	RMSE	NRMSE
Small	27s	0.96	0.48 kW	7%
Medium	231s	0.73	12.15 kW	16%

4.6.2 System identification model adaptation

SID model adaptations, as discussed earlier, are applied in the medium building case to improve the performance of the SID model. The details and results are reported in this section.

4.6.2.1 System excitation: variable, frequency and injection interval modification

In the original SID model, the excitation signals for temperature setpoint and schedule in each zone are excited together. However, the building energy system in the medium building is more complex and more nonlinear, as the results shown in Figure 4-5b, where the nonlinearity indexes are much smaller than those for the small building. Therefore, it needs more intense excitation to get enough training data for the SID model development. In this case, each zone temperature is excited individually in the medium building. For medium building, the excitation signal frequency for the adapted SID model is recalculated based on its response time (Eq. 3.5). As indicated by the system response time test results (shown in Table 4-5), the system response time of the small building is within 30 minutes in all of the testing cases (except the one of turning off HVAC system at 18 ºC). The response time of the medium building is above 30 minutes in most of the testing cases. Therefore, the 30-min excitation interval in the original system excitation plan for the medium building is needed to be extended to 60 minutes to allow enough response time for the tested system.

4.6.2.2 System identification: sampling window length modification

Sampling window length is a very important factor in the SID model, which is the length of a sampling segment. The sampling window length is determined by checking whether the Markov

parameters converge at the end of sampling window. From the Markov parameter plots in Figure 4-8, the Markov parameters for the medium building do not converge to 0 at the end of 6-hour sampling window, but they converge to 0 at around 3th hour for all the inputs of the small building. In order to determine the sampling length, a parametric test is conducted for the sampling window length as 4 hours 8 hours, 12 hours for the medium building case. The Makarov parameters converge best at end of 12-hour sampling length. Therefore, the sampling window length is adapted to 3 hours for the small building and to 12 hours for the medium building.

4.6.3 Adapted System identification model results

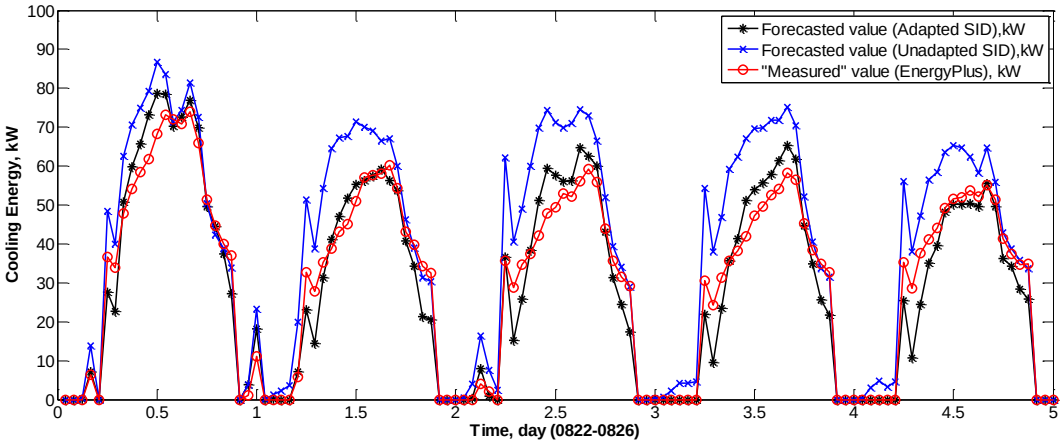
The adapted SID model is developed using the same training data as described before and is validated using the same testing data as for the original SID model. The performance comparison is tabulated in Table 4-7. It is clearly showed that the forecasting accuracy of the adapted SID model improved from 0.91 to 0.96 for the small building, and from 0.73 to 0.94 for the medium building.

Table 4-7. Building SID model performance comparison

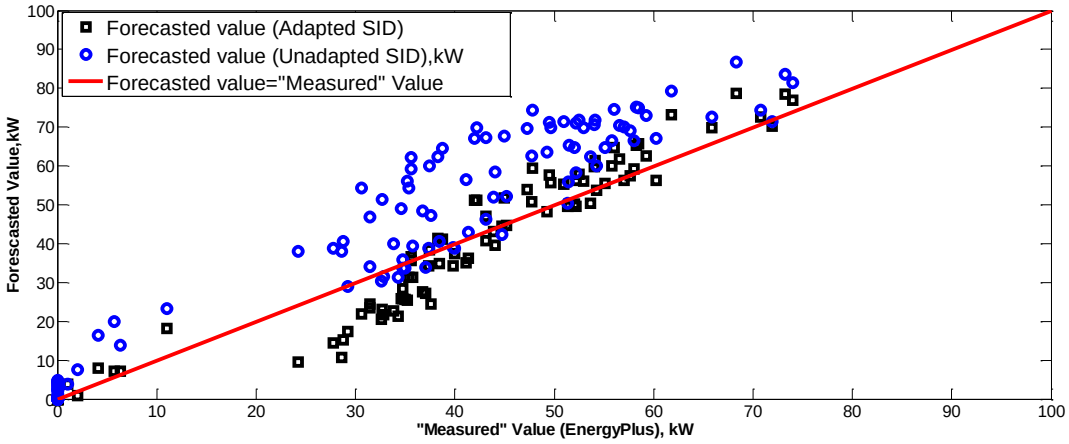
Model	Excitation Interval	Sample length	R^2	RMSE	NRMSE
Small unmodified	30 min	6 h	0.91	0.66 kW	9%
Small modified	30 min	12 h	0.96	0.48 kW	7%
Medium unmodified	30 min	6 h	0.73	12.15 kW	16%
Medium modified	60 min	3 h	0.94	5.9 kW	8%

The detailed cooling energy forecasting results from the adapted and pre-adaption SID models are compared in Figure 4-11 for the medium building. Figure 4-11a provides the time series comparison and Figure 4-11b summarizes forecasting error for pre-adaption and adaption SID models. This figure shows that the pre-adaption SID model (blue circle) overestimates the

energy consumption during most of the forecasting period, while the adapted SID model (black circle) produces better forecasting accuracy



a. Medium building energy forecasting results



b. Medium building energy forecasting error analysis

Figure 4-11. Medium building cooling energy forecasting results comparison and error analysis

4.7 Building Energy Forecasting Model Performance Comparison

In this section, the performance of the adapted SID model is compared with four building energy forecasting methods that have been reported in the literature, i.e. RC & Chiller model [110, 112], SVR [86], ANN model [82], and state space model (N4SID) [184]. The small commercial building described above is used as the testing subject. In the following comparison study, all of

the models (except adapted SID model) are provided with a 14 day (August 1st to August 14th in TMY data) training data. The training data for the adapted SID model under system active excitation lasts from August 1st to August 7th. Testing data generated from July 1st to September 30th are used for model performance comparisons. Besides the three evaluation indexes used before, forecasting extendibility and uncertainty are added into this comparison study.

4.7.1 Other models: RC & Chiller, SVR, ANN, N4SID

4.7.1.1 RC with chiller model

RC model has been applied in a lot of studies for building energy estimation with good accuracy and computational efficiency.

In the comparison study, the RC model is developed by Braun and Chaturvedi [110] and the chiller model is based on the model reported in [185]. The chiller is modeled as a third-order polynomial equation of outdoor temperature and chilled water supply temperature:

$$P_{ch} = a_1 \cdot Q_{rated} \cdot (a_2 + a_3 T + a_4 T^2 + a_5 T^3 + b_1 Q + b_2 Q^2 + b_3 Q^3 + b_4 TQ) \quad \text{Eq. 4.16}$$

$$T = (T_{wb} - T_{chws})/(T_{wb} - T_{chws})_{rated} \quad \text{Eq. 4.17}$$

$$Q = Q_{ch}/Q_{ch,rated} \quad \text{Eq. 4.18}$$

Where, $Q_{ch,rated}$ (ton) is the rated chiller capacity; T_{chws} (°F) is chiller water supply temperature, which is assumed as 35 °F in this research; T_{wb} (°F) is ambient wet bulb temperature; Q_{ch} (ton) is the cooling load

All of the parameters in the RC with Chiller model are determined through pattern searching based optimization methods in MATLAB [168]. The objective function of this optimization is described as:

$$J(Rs, Cs, ps) = \sqrt{\frac{\sum_{j=1}^{N} \left(Q_{RC,j} - Q_{Act,j}\right)^2}{N-1}} \quad \text{Eq. 4.19}$$

Where, Q_{RC} and Q_{Act} are the cooling energy consumption from RC with chiller model and EnergyPlus model, respectively; N is the total time step of this whole simulation. R_s, C_s, and p_s are parameters in the RC & Chiller model.

4.7.1.2 SVR model

The support vector regression SVR model is developed based on support vector machine classification. SVR approximate the function using:

$$f(x) = \omega \cdot \phi(x) + b \quad \text{Eq. 4.20}$$

where $f(x)$ represents the feature spaces which are mapped from the input space x. The coefficients ω and b can be estimated by solving an optimization problem:

$$J = min\ \frac{1}{2}\|w\|^2 + K(x, \hat{x}) \quad \text{Eq. 4.21}$$

$$\text{Subject to} \quad \begin{cases} y_i - f(x_i) \le \varepsilon \\ f(x_i) - y_i \le \varepsilon \end{cases}$$

The first term $\|w\|^2$ is called regularized term. The second term $K(x, \hat{x})$ is a function kernel term, which measures the error. In this study, radial basis function kernel (Eq. 4.22) was used in LibSVM [186].

$$K(x, x') = exp\left(-\frac{\|x - \hat{x}\|^2}{2\sigma^2}\right) \quad \text{Eq. 4.22}$$

Where, x and $\hat{x}$ are real and forecasted energy consumption, respectively; σ is user-defined parameter.

4.7.1.3 ANN model

ANN is a supervised learning model inspired from biological neural networks, and is widely used to estimate functions. ANN generally consists of a network of notes (neurons) and connections. The connections have different weights that are required to be tuned in the training

process. As introduced in section 1.2.2, ANN has also been successfully applied to forecast the performance of buildings and building energy systems [104, 187].

In this study, an ANN model with 10 sublayers is developed using a Matlab Neural Network Toolbox [188]. The inputs and output of the SVR and ANN models are the same as the SID model, described previously.

4.7.1.4 State spate model

In this comparison study, the N4SID model in MATLAB system identification toolbox is used to develop a 5-order state space model to forecast the building cooling energy consumption [189]. N4SID has been used and recommended in a lot of literature for the good performance in building energy forecasting, such as [190, 191]. The N4SID model is able to identify the system order and model parameters (A, B, C and D) for a state space model:

$$\frac{dx}{dt} = A\,x + B\,u \qquad \text{Eq. 4.23}$$

$$y = C\,x + D\,u \qquad \text{Eq. 4.24}$$

where, x is the state vector, which cantinas all the forecasting state variables, u is the control vector containing the control variables, y is the measurement vector. In this comparison study, x is defined as the system inputs: T_{out}, Q_{dir}, Q_{dif}, and V_{oa}. u is the control variables: $T_{zone,\,stp,\,i,}$ and $R_{in,i}$.

4.7.2 Comparison Results

In this section, the performance of the developed SID model is compared against 4 other common system forecasting models. The forecasting accuracy, speed, extendibility and uncertainty comparison results are presented in the following sections.

4.7.2.1 Accuracy and speed

Firstly, all these 5 models developed upon the training data described above are used to forecast the cooling energy consumption from August 22 to August 28. The energy forecasting results from these 5 different models are compared with the EnergyPlus simulation results in Figure 4-12. The forecasting accuracy and speed comparison is summarized in Table 4-8. The results show that the SID model achieved the highest forecasting accuracy, where R^2, RMSE and NRMSE are 0.96, 0.48 kW and 7%, respectively.

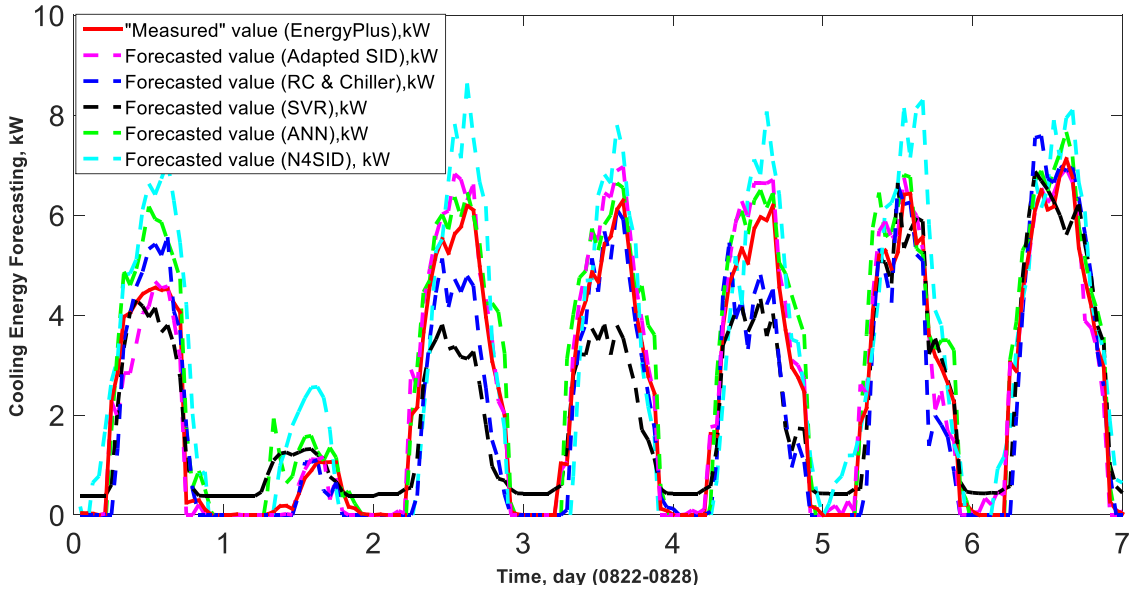


Figure 4-12. Small building cooling energy forecasting case study results

Table 4-8. Cooling energy forecasting model performance

Model	L/NL	Excitation	Training Period	Testing Period	Calculation Time	R^2	RMSE	NRMSE
SID	L	Yes	0801-0807	0822-0828	21s	**0.96**	**0.48 kW**	**7%**
N4SID	L	No	0801-0814	0822-0828	2.88s	0.89	1.11 kW	15.6%
RC & Chiller	NL	No	0801-0814	0822-0828	451s	0.87	0.83 kW	11.70%
SVR	NL	No	0801-0814	0822-0828	**0.02s**	0.69	1.04 kW	14.59%
ANN	NL	No	0801-0814	0822-0828	2.78s	0.93	0.68 kW	9.60%

Even though SID model requires longer calculation time than SVR and ANN, there is only 0.004 second energy forecasting time and almost all the time are used in the model training period. Fortunately, in any model based control method, the energy forecasting model just need to be trained once off line within certain forecasting period. Therefore the SID model speed is still acceptable. Besides the three index discussed before, another evaluation factor, fractional bias (FB) (Eq. 4.25), is also used to compare the model performances.

$$FB = 2 \cdot \frac{x_{ture} - x_{predict}}{x_{ture} + x_{predict}} \qquad \text{Eq. 4.25}$$

Where, x_{ture} is the "measured" energy consumption, $x_{predict}$ is the predicted energy consumption, and FB is the factional bias. The FB will have a value of 0 when x_{ture} and $x_{predict}$ perfectly and will tend towards −2 or 2 as these quantities differ by greater magnitudes. The FB evaluation results are tabulated in Table 4-9. As this table shown, in five days of this week, the FB of SID model is closest to 0. Even though average FB of RC & chiller model (0.21) is closest to 0, the variation in these 7 days is very large from positive to negative. Therefore, in this fractional bias evaluation, the SID model is superior to other models.

Table 4-9. Energy forecasting fractional bias

RMSE	**Day 1**	**Day 2**	**Day 3**	**Day 4**	**Day 5**	**Day 6**	**Day 7**	**Average**
SID	**-0.25**	0.11	**-0.43**	**-0.37**	**-0.41**	**-0.42**	-0.39	-0.31
N4SID	-1.35	-2.00	-1.07	-0.85	-0.91	-1.53	-0.77	-1.21
RC & chiller	-0.87	**0.09**	1.56	0.13	1.75	-0.51	-0.65	0.21
SVR	0.58	-1.31	2.00	1.70	1.68	-0.47	**0.17**	0.64
ANN	-1.11	-1.21	-1.66	-0.91	-0.97	-1.05	-0.30	-1.03

4.7.2.2 Extendibility

In order to test and compare the extendibility of all five models, two folds of testing studies have been designed. The first one is to test the model performance with different weather conditions, and the other one is to test the model performance with different operation strategies, such as temperature setpoints.

In the first direction, three scenarios are designed using the same training period as described previously. The testing periods for the three scenarios are designed so that Scenario 1 represents the situation when weather conditions during the testing period are similar to those during the training period. Scenario 2 and Scenario 3 are two months before and after the training period when weather conditions are, sometimes, out of range of the weather conditions during the training period. Table 4-10 summarizes the temperature during training and testing periods for the three scenarios. Other than the temperature in all training and testing period in all these three scenarios there are sunny and cloudy days with high and low solar radiation.

Table 4-10. Model extendibility testing condition

Scenario	Testing Period	Training Temperature Range ºC	Testing Temperature Range ºC
1	0801-0831	20.1-35.6	17.6-31.1
2	0701-0730	20.1-35.6	12.0-37.3
3	0901-0930	20.1-35.6	6.7-34.0

The five models are applied in all these three scenarios to forecast the cooling energy consumption. The forecasting result comparison is summarized in Table 4-11. For Scenario 1, in which the weather conditions during the testing period is more similar to the weather condition during training period, the ANN model achieves the highest R^2 (0.93), followed by the adapted SID model, which achieved a similar R^2 (0.92). For the same scenario, the adapted SID model achieves the lowest NRMSE (7.4%), followed by the SVR and ANN models. However, in

Scenarios 2 and 3, in which weather conditions re more different to those in the training period, the adapted SID model performs much better than other four models and achieves the highest R^2 and lowest NRMSE. Therefore, the adapted SID model performs much better when the weather conditions during the forecasting period are quite different from those during the training period and thus has much better extendibility than other four models.

Table 4-11. Cooling energy forecasting model extendibility

Scenario	Forecasting accuracy, R^2					Forecasting accuracy, NRMSE				
	Adapted SID	N4SID	RC & Chiller	SVR	ANN	Adapted SID	N4 SID	RC & Chiller	SVR	ANN
1	0.92	0.83	0.77	0.90	**0.93**	**7.4%**	11.6%	16.3%	9.6%	9.8%
2	**0.86**	0.67	0.81	0.56	0.79	**9.6%**	18.4%	13.1%	28.2%	15.3%
3	**0.89**	0.55	0.75	0.88	0.73	**7.9%**	33.9%	17.8%	10.6%	18.3%

In order to test the model extendibility against different operation signals, a temperature setting schedule from an existing demand response study is used here. Similar to the previous studies, the SID model is training using building operation data under system excitation, while all the other models are based on the building operation data under regular strategies. This comparison study is also conducted in the small commercial building described in section 3.5.1. The temperature setpoints in the training periods are: 26.7 ºC from 0 am to 6 am, 24 ºC from 6 am to 6 pm, and 26.7 ºC from 6 pm to 12 am. Then these five different models are used to forecast the energy consumption for five weekdays from August 24 to August 28. The temperature setpoints in these five days are: 32 ºC from 12 am to 4 am, 18 ºC from 4 am to 6 am, 24 ºC from 6 am to 6 pm and 32 ºC from 6 pm to 12 am, as tabulated in Table 4-12.

The cooling energy forecasting results from all these 5 models are plotted in Figure 4-13. It roughly shows that the “Adapted SID” model and “RC & Chiller” model have the better accuracy.

Table 4-12. Training and forecasting period temperature setpoint

Time	Cooling Setpoint, ºC (ºF)	
	Training	DR forecasting
0 - 4 am	26.7 (80.1)	32 (89.6)
4-6 am	26.7 (80.1)	18 (64.4)
6 am – 6 pm	24 (75.2)	24 (75.2)
6 pm – 12 am	26.7 (80.1)	32 (89.6)

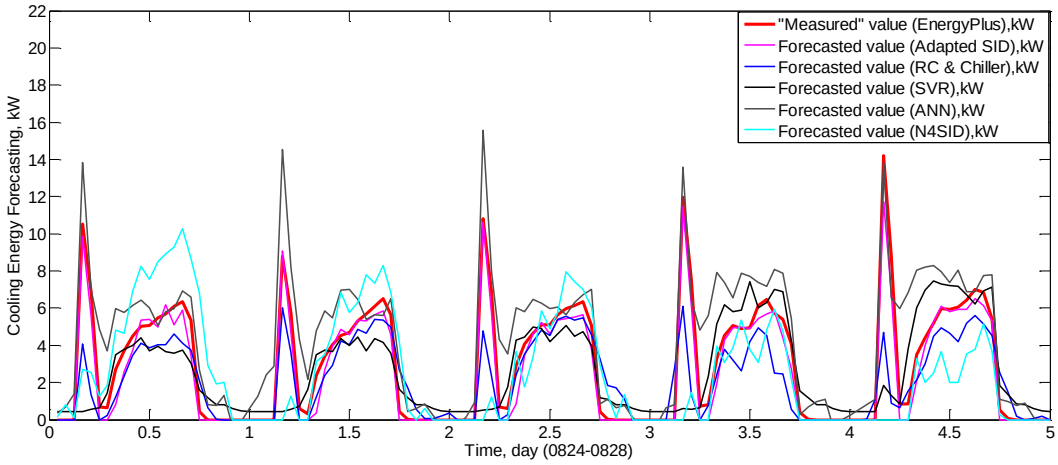


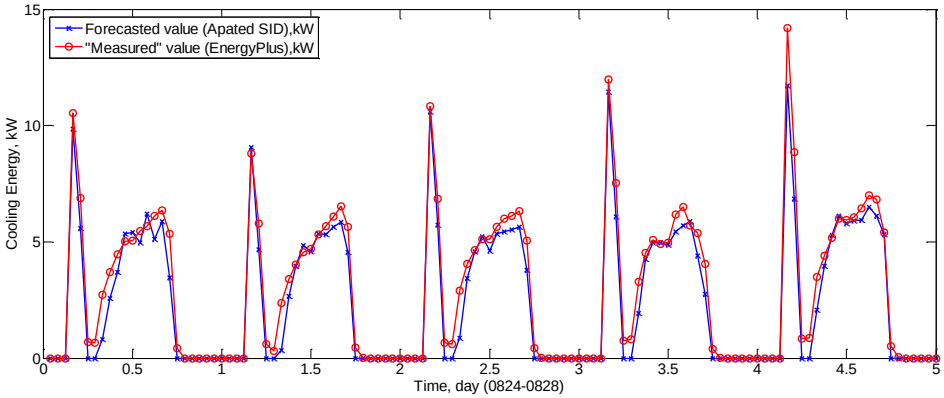
Figure 4-13. Small building cooling energy forecasting under DR operation

The detailed energy forecasting accuracy statistics are tabulated in Table 4-13. The adapted SID and RC & Chiller model achieved the highest R^2 above 0.8, followed by the SVR model of 0.73. For the forecasting error, obviously, adapted SID and RC & Chiller model have the lowest RMSE and NRMSE.

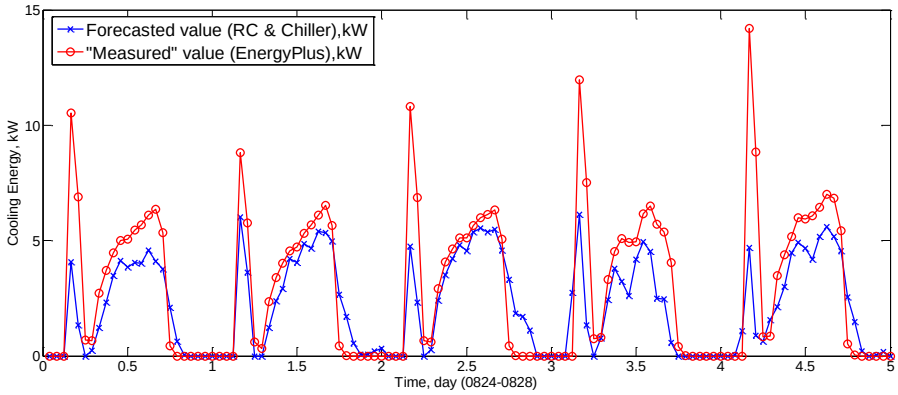
Table 4-13. Energy forecating accuracy under DR opertion

	R^2	RMSE	NRMSE
Adapted SID	**0.86**	**1.15**	**8.1%**
N4SID	0.44	2.62	18.0%
RC & Chiller	**0.81**	**1.32**	**9.0%**
SVR	0.73	1.89	13.4%
ANN	0.57	2.35	17.3%

In order to examine the forecasting result clearly, the forecasting results from these two models are plotted in Figure 4-14. In Figure 4-14a, the SID model mostly captured the trend of the "measured" value, even in the pre-cooling period in the early morning. For the results of RC & chiller (Figure 4-14b), it cannot capture the sudden increase of the cooling energy due to the precooling very well, but it is also able to forecast the energy consumption at other times.



a. SID model forecasting results



b. RC& Chiller model forecasting results

Figure 4-14. Small building cooling energy forecasting under DR operation

4.7.2.3 Model Uncertainty

A Monte Carlo (MC) simulation is conducted to analyze the noise impact on model accuracy. During a MC simulation, the following process is used:

1. Initialize Monte Carlo simulation by defining input noise distributions (+/- 5%) and adding the noise to the measurement (simulation results from the virtual building);
2. Perform MC: for i=1...N (in this study, N is chosen as 1000)
    - Sample noise values from defined distributions
    - Run each model for energy forecasting
    - Calculate MC output (daily energy consumption, kWh) from each model
3. Analyze the performance of each model

In this MC simulation, 5% Gaussian distributed random white noise is added into each input variable, and the N is 1,000. Daily energy consumption is chosen as the output of the MC simulation. The Boxplots of the MC simulation output during the testing period (August 22nd to August 28th) are shown in Figure 4-15. In this figure, each box shows the 5% and 95% percentile of the energy forecasting distribution, the middle lines show the mean of the daily energy consumption (kWh) during the MC simulation, and the red dash lines are the cooling energy consumption under noise-free condition from EnergyPlus.

As Figure 4-15 shown, the SVR model and ANN models have lower uncertainty which is reflected by the narrow boxes, their accuracy is much worse than that of the adapted SID model and RC & chiller model, by observing the differences between the medium lines and red dash lines. It means that SID model and RC & chiller model are more sensitive to the input noise, while they still can maintain better accuracy than SVR model and ANN model.

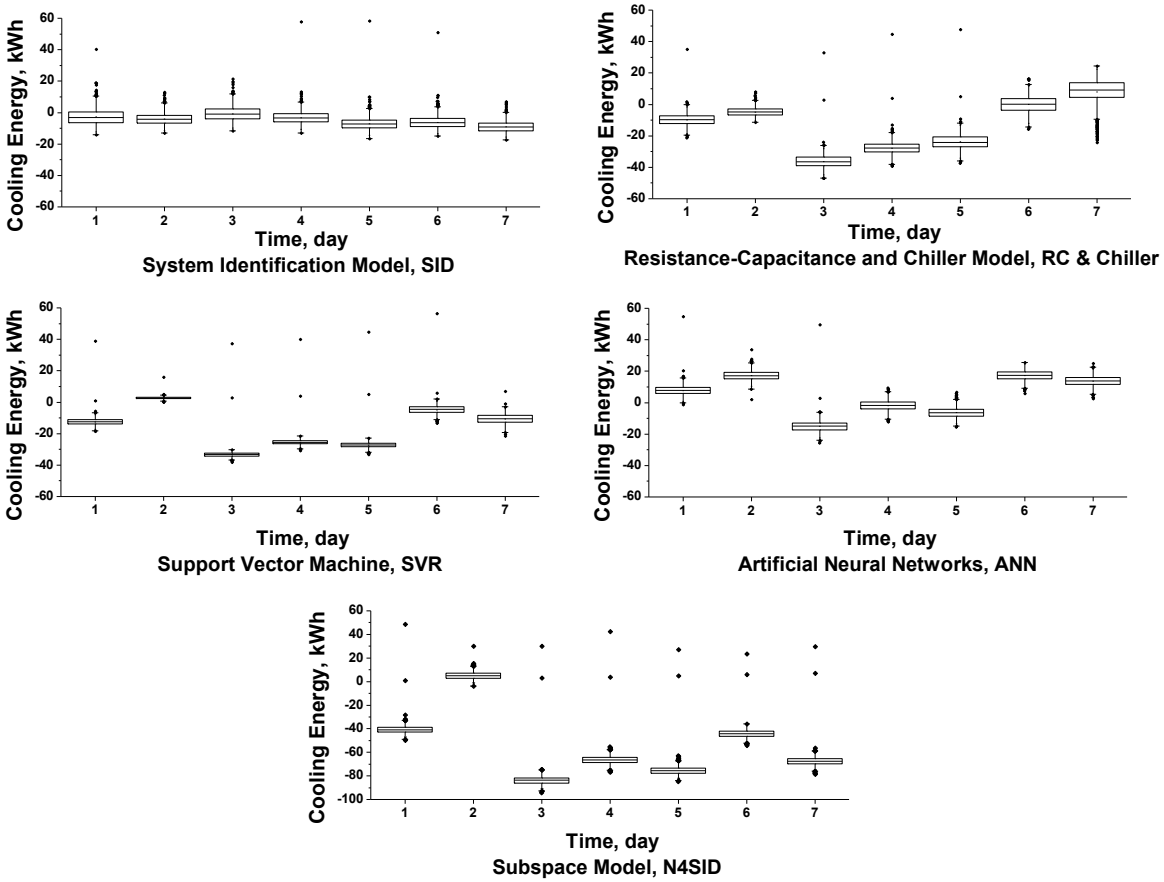


Figure 4-15. Boxplots of Monte Carlo daily energy consumption simulation results

The energy forecasting accuracy in the MC simulation of each model are summarized in Table 4-14, which shows SID model maintains the lowest RMSE in 4 days (Day 1, 3,), while RC & chiller model achieves the highest accuracy in Day 2. By examining the weather condition, the temperature of Day 2 is much lower than the temperature of other days, and the cooling energy consumption in Day 2 is also much lower than that in the other days. Because the RC & chiller model includes the physics theory into the model, it is able to maintain the higher accuracy when the weather conditions suddenly change. Therefore, RC & chiller model achieved the highest energy forecasting accuracy at Day 2 in the MC simulation. In Day 4 and 5, the ANN model achieved the lowest RMSE, while the results of SID model are just slightly larger than those of

ANN model. On the other hand, the MC simulation results of N4SID model have the highest RMSE in 6 days (Day 1, 3-7) during this 7-day simulation study, and in Day 2 the RMSE of N4SID simulation result is slightly better than that of ANN model. As a result, it is concluded that SID model has the lowest uncertainty and N4SID model has the highest uncertainty in MC simulation study. By comparing the average RMSE of these 7 days, SID model achieved the lowest error, and is 47.7 % less than the second lowest model.

Table 4-14. Energy forecasting accuracy in MC simulation, RMSE (kWh)

RMSE	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7	Average
SID	**6.08**	3.88	**4.80**	4.82	6.51	**4.83**	**10.55**	**5.9**
N4SID	19.5	21.9	39.9	35.7	38.7	38.9	43.0	33.9
RC & chiller	10.98	**2.65**	36.41	27.65	22.53	8.00	10.89	17.0
SVR	13.17	7.40	33.25	25.25	25.57	11.15	12.09	12.8
ANN	7.87	22.12	15.26	**3.74**	**5.58**	11.55	13.23	11.3

4.8 System Identification Method Application in Large Building

After this system identification methodology for energy forecasting has been applied successfully in the small and medium office buildings, this method is then applied in a large size office building. The "Large office" DOE commercial reference building has been used here for the verification. This building has 12 stories and a basement. The total area is 46,320 m^2, and the window-to-wall ratio is 38%. The overall U-factor of its single pane windows is 3.2 W/m^2K with a solar heat gain factor of 0.36. The overall U-factor of the external envelopes is 0.86 W/m^2K. The U-factor and solar absorptivity value of the roof are 0.36 W/m^2K and 0.7, respectively. The HVAC systems used in this building are signal duct VAV with reheat systems.

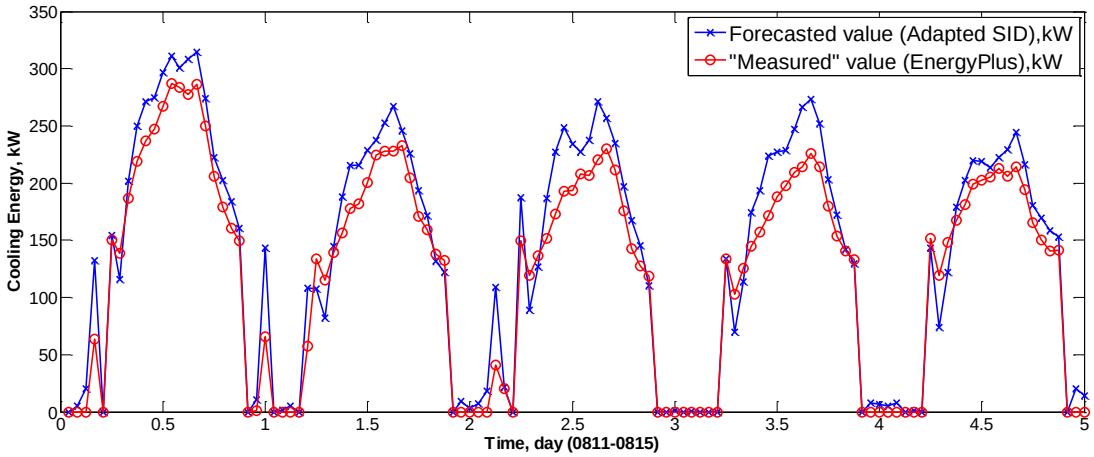
Following the procedures of system characteristics test, system excitation and system identification, the SID model has been developed to forecast the energy consumption. In this

study for this 12-story building, only 3 different SID model has been trained individually for bottom floor (first floor), medium floor (2 – 11 floors), and top floor (12 floor). This is because the structure and systems for 2-11 floors are very similar. The only difference are the different operation strategies, such as temperature setpoints and equipment schedules, while the SID model developed in this study is capable of handling different operation schemes. The first and top floor, however, have more different disturbance, such as heat transfer from the basement, solar heat gains from the roof. Therefore, individual SID model has been developed for the first and top floor. The energy forecasting results are summarized in Table 4-15, where the energy forecasting accuracy can reach to 0.84 (R^2) and the forecasting error is within 12.1% (NRMSE).

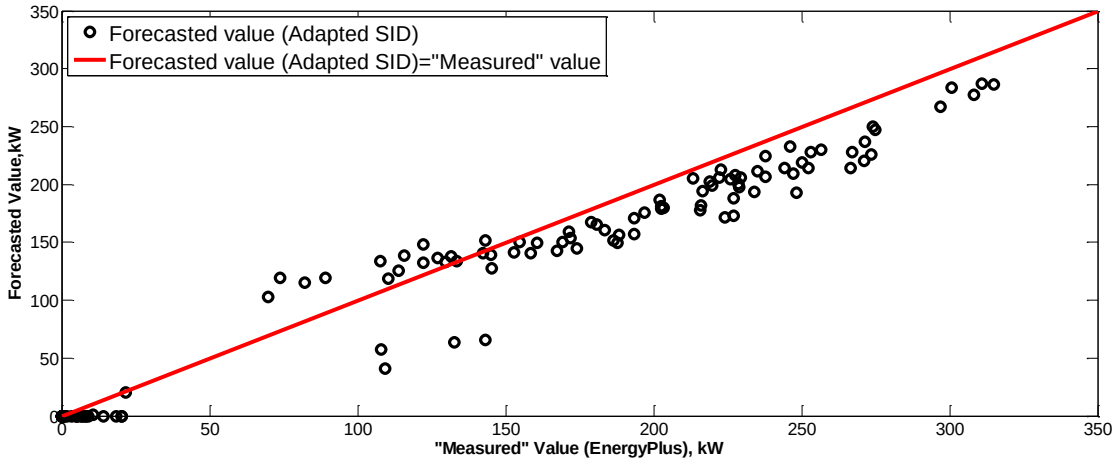
Table 4-15. Energy forecasting results for large office building

	Period	Temperature Range, ºC	R^2	NRMSE	Calculation Time, S
Training	0801-0807	20.1-35.6	0.95	5.32%	165.72
Forecasting	0811-0815	18.0-35.0	0.91	9.0%	0.11
Forecasting	0822-0826	20.0-35.1	0.84	12.10%	0.13

The detailed forecasting results at each hour are plotted in Figure 4-16. As it shown, the SID model over-estimated the energy consumption at early afternoon, which has also been identified in the small and medium building case.



a. Large building energy forecasting results



b. Large building energy forecasting error

Figure 4-16. Large office energy forecasting results

4.9 Conclusion and Future Work

This section extended a previous study and presented methods to test a building energy system's nonlinearity and response time and the system identification model adaptation based on the characteristic test results. The proposed methods are applied to two different buildings with varying characteristics. It is found that three key system identification model parameters, i.e. the system excitation signal generation frequency, injection interval, and sampling length, need to be adapted to ensure the model forecasting accuracy for different building energy systems. The energy forecasting results show that the adapted SID model accuracy is maintained above 90% for the small building case and is improved from 73% to 94% for the medium building case. The adapted SID model is also compared against 4 reported building energy forecasting models: N4SID model, RC & chiller model, SVR model, and ANN model in the model accuracy, speed, extendibility and uncertainty. Building operation data with and without sensor noise are generated for the comparison study. As the model comparison study results show, SID model has the capability to achieve higher accuracy and extendibility under both of the noise-free and noisy conditions. Beyond the model adaptation and comparison work, this system identification

methodology for energy forecasting has then been applied in a large office building. The energy forecasting accuracy achieved is also over 80%. Future efforts are needed to improve the excitation strategy and to examine the developed methodologies in real world conditions.

5. CHAPTER 5 Building Cluster Energy Forecasting: Other Modeing Approaches

5.1 Background and Motivation

Although the major effort in this study has been developing on-line energy forecasting model using system identification technique and the feasibility study of system identification methodology has proven to be a promising approach, grey box modeling approaches, especially when used on alternative energy and energy storage devices, are studied in this chapter. Due to the relatively simpler operation ranges, it is much easier to collect enough operational data for grey box model development for alternative energy and energy storage devices. Therefore, in this chapter, the most promising gray box models for whole building system, alternative energy generation device and energy storage devices are discussed first. Simulation operational data generated from the cluster emulator developed in Chapter 2 is used to develop and examine these models.

5.1.1 Grey Model for Whole Building Modeling

The common grey models for whole building modeling have been reviewed in Section 1.2.3. In this Chapter, RC model has been chosen to estimate the building energy consumption.

5.1.2 Ice Tank Thermal Energy Storage Modeling

Besides battery electrical storage system, ice tank thermal energy storage system is another building energy management equipment. Building operator can also use ice tank to shave the high electricity demand for cooling load during peak hours associated with real time electricity price. In order to fully utilize the peak demand shaving, an accurate and efficient ice tank thermal storage model is needed.

West and Braun presented and validated an empirical model for simulating an ice tank [192]. The basic physical heat transfer of the ice tank model is modeled by Eq. 5.1 to Eq. 5.3:

$$X_k = X_{k-1} + \frac{u_k \Delta t}{Ca_s} \quad \text{Eq. 5.1}$$

$$u = \varepsilon_c m_f c_f (T_s - T_{f,i}) \quad \text{Eq. 5.2}$$

$$\varepsilon = \frac{T_{f,i} - T_{f,o}}{T_{f,i} - T_s} \quad \text{Eq. 5.3}$$

Where, X is the ice storage state (state of charge), Δt $(h)$is the charging/discharging time, u (w) is heat transfer rate (charging/discharging rate), Ca (wh) is the storage capacity, ε is heat transfer effectiveness, m_f is second fluid mass flow rate, $c_f(\frac{\mathbf{J}}{\mathbf{kgK}})$ is the specific heat of second fluid, T_s, $T_{f,i}$, and $T_{f,o}$ (k) are melting/freezing temperature, second fluid inlet, and outlet temperature, respectively. Empirical studies have been conducted to determine the heat transfer effectiveness, ε, at different states of charge (SOC) and other situations. The overall predicted charging and discharging effectiveness were within about 4% of real field measurements.

Ihm et al. [37] developed an ice tank TES module and integrated it with EnergyPlus. This model is based on a previous study in [192]. Three different operation modes, namely dormant mode, charging mode and discharging mode, were modeled and validated. When the TES system is in dormant mode, the charging/discharging rate, u, is zero. In charging mode, TES SOC is calculated in Eq. 2.2. During the discharging period, the TES system provides cooling to meet the cooling demand from the supply side. The TES water flow rate, $\dot{m}_{ice}$, is adjusted based on the load request, $\dot{Q}_{ice}$, and inlet chilled water temperature, T_{inlet} , as Eq. 2.3:

Based on one of benchmark building EnergyPlus models provided in [15], Sehar et al. [40] analyzed the chiller energy consumption of conventional non-storage and ice storage cooling systems for large and medium-sized office buildings in different climate zones. The ice tank

model in this paper is based on EnergyPlus ice storage model [37]. The impact of ice storage on the building energy consumption in different cities, such as Miami, Las Vegas, Baltimore, Seattle, Chicago, Helena, and Duluth have been analyzed and compared. From their results, they concluded that with full storage ice tank there was no energy consumption at peak hour, round 50% peak hour energy consumption reduction with storage priority operation, and 25% reduction with chiller priority operation [40]. They found out that the ice storage systems had higher chiller energy consumptions than the conventional non-storage systems due to the night dedicated chiller operation. But the ice storage system can reduce or even eliminate the chiller operation during the peak hours, which can save energy cost. Climate zones with summers having high temperatures and relative humidity ratio increase not only the building cooling load but also the chiller energy consumption by decreasing the cooling of condenser water.

Henze et al. [38, 193] developed and validated a simulation environment for the evaluation of the performance of various controls of ice storage system. Chiller-priority, constant-proportion, and storage-priority control strategies was compared to the optimal control strategy that achieves the theoretical maximum of operating cost savings. Dynamic programming based global search in peak and off peak demand domains was applied to find out the ice tank storage optimal operation schemes. A set of ice storage system's operation guidelines under different conditions were identified to improve the load-shifting performance of different control schemes. Hajiah and Krarti [194, 195] presented a novel simulation environment which can analyze the benefit of ice storage system as well as building thermal mass simultaneously to reduce the energy cost while maintaining occupant comfort. The optimal objective function, as described in Eq. 5.4, was adopted from [193].

$$C = r_{d,o}P_{max,o} + r_{d,1}P_{max,1} + \sum_{k=0}^{K} r_{e,r(k)}P(k)\Delta t \qquad \text{Eq. 5.4}$$

Where C is the total cost including energy and demand charges ($), $r_{d,o}$ is the off-peak demand charge ($/kW), $r_{d,1}$is the peak demand charge ($/kW), $P_{max,o}$ is the off-peak electricity demand, $P_{max,1}$is the peak electricity demand, r_e is the energy charge ($/kWh), k is the number of hours in each simulation period, $P(k)$ is the total power consumption at time k.

Direct search complex method was used to solve the nonlinear constrained minimization objective functions. Several other optimization methods have also been investigated to improve the effectiveness of the ice storage system. Based upon the optimal operation control schemes, round 10.8% total cost saving, including base cost and demand cost, was achieved in a laboratory located in Boulder, CO. The results of the validation analysis indicated that the simulation environment predict cost savings for optimal controls with 10% agreement when compared to the experimental measurements. Quasi-Newton method was investigated by Henze at al. [111] to optimize the building operation in response to time-of-use electricity rate by using an ice storage system and building thermal mass. They concluded that when an optimal controller was given perfect weather forecasts and when the building model used for predictive control matched the actual building, utility cost savings and on-peak electrical demand reductions were substantial. Chen et al. [196] explored the optimization of ice storage air conditioning system by using dynamic programming algorithm. They developed the chiller power consumption models as well as the ice tank storage heat transport model based on manufacture data. The initial cost and operation cost are objective functions. The ice storage tank was modeled as a heat exchanger, which discharging rate Q_{ice} was expressed as:

$$Q_{ice} = UA_{ice} \times \Delta T_{lm,ice} \quad \text{Eq. 5.5}$$

$$\Delta T_{lm,ice} = \frac{\left(T_{ice,in} - T_f\right) - \left(T_{ice,out} - T_f\right)}{ln\left(\frac{T_{ice,in} - T_f}{T_{ice,out} - T_f}\right)} \quad \text{Eq. 5.6}$$

where $T_{ice,in}$, and $T_{ice,out}$ are inlet and outlet brine temperature, and T_f is freezing temperature of water, i.e. 0 ^{o}C.

Massie [197] developed and tested a neural network-based optimal controller for an ice tank thermal energy storage system. This controller can self-learn the behavior of equipment and then determine the system operation scheme to minimize its operation cost. The controller consists of four neural networks, three of which map equipment behavior and one that acts as a global controller. The ice tank controller consists of a training and predictor network working in parallel. The training network captures the relationship between controlled variables and other ice tank performance variables, such as chilled power consumption, and ice tank charging/ discharging rate. This optimal control model does not rely on any rules or assumptions, it can find out the optimal operation solution at any pricing structure, building load profile and equipment operation conditions.

5.1.3 Solar Radiation Model and PV Panel Power Generation Modeling

A number of building energy modeling and building operation researches with short term solar radiation have been undertaken in the past twenty years, using different statistical methods, such as autoregressive moving average (ARMA) [198-200], artificial neural networks [201-203]. The general approach for PV panel power generation estimation is: (i) estimating the total solar irradiance on the tilted surface of the PV panel from overall global solar radiation on the horizontal surface, (ii) calculating absorbed solar irradiance based on the solar irradiance on the PV cell surface, and (iii) computing the PV power generation from absorbed solar irradiance. The total solar irradiance on the inclined surface (G_{tt}) can be calculated in terms of direct beam (G_{bt}), sky diffuse (G_{dt}) and ground reflected (G_{rt}) solar irradiance individually:

$$G_{tt} = G_{bt} + G_{dt} + G_{rt} \qquad \textit{Eq. 5.7}$$

Traditionally, direct beam solar radiation calculation is straight forward from beam solar radiation on a horizontal plane, incidence (θ) and solar zenith (θ_z) angles [204] (Eq. 5.8).

$$G_{bt} = G_{bh}\frac{cos\theta}{cos\theta_z} \qquad \text{Eq. 5.8}$$

Ground reflectance solar radiation is usually estimated by Eq. 5.9 [204]:

$$G_{rt} = \frac{\rho_0}{2}G_{th}(1 - cos\beta) \qquad \text{Eq. 5.9}$$

where G_{th} is the total irradiance on a horizontal plane, and $\rho_0$is ground reflectance, and β is the slope angle of the tilted surface. The isotropic sky model is the simplest model with the assumption that all diffuse solar radiation is uniformly distributed over the sky [205].

However, there are a lot of new techniques to estimate the solar radiation on PV panels, especially to estimate the diffuse solar irradiance. ARMA model, taking deterministic annual and diurnal periodicity and the variation of weather condition variables into account, was used by Yoshida and Terai to estimate the ambient air temperature, solar radiation, and absolute humidity. Each weather condition variable was decomposed into deterministic or periodic and random components. The deterministic component was modeled by a Fourier series and the other one was modeled as an ARMA model [199]. Zhou et al. [113] integrated their on-line grey box building energy model with on-line air temperature, relative humidity and solar radiation prediction models to increase the next day hourly building load forecasting accuracy. The regressive empirical solar radiation forecasting model used the forecasted cloud amount, extreme temperatures from the weather stations (Eq. 5.10).

$$I_{global} = I_0\left(a_1 + a_2\sqrt{T_{max} - T_{min}} + a_3\sqrt{1 - \frac{C}{8}}\right) \quad \text{Eq. 5.10}$$

Where X is the forecasting variable, including temperature and relative humidity, a and b are coefficients determined by regression model. In Eq. 5.10 a_1, a_2 and a_3 are regression coefficients obtained from recursive least square algorithm, I_{global} is hourly global solar radiation, T_{max} and T_{min} are daily maximum and minimum temperature, C is average daily cloud coverage, and I_0 is hourly extraterrestrial solar radiation. Less than 10% variance between observations and prediction was achieved through this reported model. Coskun et al. [206] proposed and demonstrated a probability density frequency based method for estimating solar radiation distribution. In this study, a case study using 15 years recorded actual global solar irradiation data was conducted to analyze its influence on the performance of solar collectors. ASHRAE [183] provided models for hourly global radiation(I), and hourly diffuse radiation (I_d) on the horizontal surface on a clear day as follows:

$$I = I_N cos\theta_z + I_d \quad \text{Eq. 5.11}$$

$$I_N = Aexp(-B/cos\theta_z) \quad \text{Eq. 5.12}$$

$$I_d = CI_N \quad \text{Eq. 5.13}$$

where θ_z is solar zenith angle. Model parameters A, B and C are given in a lookup table.

Another diffuse solar radiation estimation model was developed and published in [207], where models for daily values of diffuse solar radiation were developed from (i) diffuse fraction or cloudiness index as a function of clearness index, (ii) diffuse fraction or cloudiness index as a function of relative sunshine duration or sunshine fraction, (iii) diffuse coefficient as a function of clearness index, and (iv) diffuse coefficient as a function of sunshine duration or sunshine fraction. Spatial-temporal covariance structures and time-forward kriging approach has been

applied by Yang et al. [208]. The methodology in this study lines in two transformations of the real weather data: (i) applying spatial-temporal covariance transformation to obtain spatial stationarity, (ii) reducing the residual sum-of-squares by fitting exponential correlation functions. Boland et al. [209] constructed a logistic model for direct normal solar radiation using solar radiation data from multiple locations, and the performance of this logistic model was then compared with that of Perez model, which turned out that the logistic model performed arguably better that Perez model [210]. Finally, Boland–Ridley–Lauret model [211] was used to obtain hourly diffuse radiation from direct normal solar radiation. Olmo et al. developed a inclined surface solar irradiation (G_β) model based on solar irradiation on horizontal surface (G) with incidence (θ) and solar zenith (θ_z) angles [212].

$$G_\beta = G\varphi_o \qquad \text{Eq. 5.14}$$

$$\varphi_o = \exp[-K_t(\theta^2 - \theta_z^2)] \qquad \text{Eq. 5.15}$$

Another recent study developed a solar 3D urban model to calculate the solar energy potential in building roofs and facades [213]. In this study, a digital surface model was built and a solar radiation model was also developed based on climatic observation data. Direct and diffuse solar radiation was then obtained for roofs and facades with considering the shadow effect. The results from this study confirmed that the annual irradiation on vertical facades is lower than that on roofs. A review paper about the global solar radiation models has been published by Bakirci [214], where 60 models for relating global radiation to sunshine hours, relative humidity ratio, temperature, etc. identified by the author were reviewed and discussed. This paper concluded that the most commonly used parameter for global solar radiation estimation is sunshine duration which is widely available. Khalil and Shaffie [215]compared different models for estimation of total solar radiation on horizontal surface, empirical correlations for global solar energy models on inclined surface, and estimation models for diffuse solar energy models on inclined surface,

such as Badescu model [216], Tian et al. model [217], Skartveit and Olseth model [218], Steven and Unswoth model [219], and Perez model [220].

PV panel is a common on-site energy generation device. Lots of existing studies are focusing on PV panel modeling and prediction to reduce energy consumption and cost. Sandia Nation Laboratories (SNL) provided a detailed description of the PV module model in [221]. The output current of the PV array can be expressed as Eq. 5.16:

$$I_P = MI_l - MI_o\left[exp\left(\frac{q\left(NV + \frac{I_p R_s N}{M}\right)}{NAKT_p} - 1\right)\right] - \left[\frac{NV + \frac{I_p R_s N}{M}}{\frac{NR_{sh}}{M}}\right] \quad \text{Eq. 5.16}$$

where I_p = output current of panel (A), I_l = light generated current per module (A), *Io*= reverse saturation current per module (A), M = number of module strings in parallel, N = number of modules in each series string, V = terminal voltage for module (V), *Rs*= diode series resistance per module (ohms), *Rsh* = diode shunt resistance per module (ohms), q = electric charge (16–19 C), k = the Boltzmann constant (13.8–23 J/K), A = diode ideality factor for the module, and *Tp* = cell temperature (K).

The National Institute of Standards and Technology (NIST) created a building integrated PV test bed to collect experimental data to improve simulation models [222]. Twelve-month's experiment data was collected and compared with the simulation results of SNL's model. The agreement of annual energy output predictions from SNL model and NIST test bed is within 7%. Lu et al. [130] developed a simple, practical model to describe the characteristics of power output of PV modules (Eq. 5.17). This model represented the *I-V* characteristics of PV modules based on the equivalent circuits of solar cells. The power output model was developed upon this *I-V* characteristics model.

$$I = I_1 - I_{01}\left[exp\left(\frac{V + R_s I}{V_t}\right) - 1\right] - I_{0m}\left[exp\left(\frac{V + R_s I}{mV_t}\right) - \frac{V + R_s I}{R_p}\right] \quad \text{Eq. 5.17}$$

Where I is the current generated by the solar cell, I_1 is the light-generated current, I_{01} is the reverse-saturation current of ideal diode, V is the voltage generated by the solar cell, R_s is the series resistance of the solar cell, V_t is thermal voltage depending on the cell temperature, I_{0m} is the reverse-saturation current of the non-ideal diode, zero in ideal case, R_p is the leakage resistor, and m is a constant.

Jones and underwood developed and validated a series of PV module electrical models based on the PV fill factor method taking solar radiation and temperature characteristics into account [223]. The PV power output from array can be estimated from fill factor (*FF*), short circuit current under stander temperature and irradiance (I_{sco}), solar irradiance on a tilted panel (G), standard irradiance (G_0), open circuit voltage (V_{oc}), module temperature (T_m), standard module temperature (T_0), number of PV modules (N_m), and the inverter efficiency (η_{in}):

$$P_{array} = FF\left(I_{sco}\frac{G}{G_0}\right)\left(V_{oc}\frac{ln(k_1 G)}{ln(k_1 G_0)}\frac{T_0}{T_m}\right)N_m\eta_{in} \quad \text{Eq. 5.18}$$

This model was validated against a real building-integrated 39.5kW PV array operation data over a period of one year. ANN method was applied by Mellit et al. for a PV generator model, a battery model, and a PV regulator model [93]. Levenberg–Marquardt algorithm and infinite impulse response filter were used to accelerate the calculation speed. The correlation coefficient between predicted and observed value varies from 90% to 96% for each estimated signal. Kim et al. [224] presented a grid-connected PV system model for transient analysis. The modeling and simulation in this study were realized in PSCAD/EMTDC, a widely used power system simulation tool. This model is capable to analyze the dynamic behavior of the power system, such as the controller operation, and anti-islanding performance. Hernandez et al. [225] presented a

methodology to predict the behavior of a gird-connected PV system from measurements of solar radiation and ambient temperature in a statistically reliable way. This methodology allows determining the probability density functions representing statistically the behavior of PV system under the uncertainty of weather conditions. The solar resource characteristics and the photovoltaic system models are integrated into a non-deterministic approach using the stochastic Monte Carlo method. The accuracy of this model depends mainly on the quality and availability of on-site meteorological data. Henze and Dodier developed a simple PV power generation model and investigated an adaptive optimal control of a grid-independent PV system consisting of PV panel, electricity storage and a building load [132]. Q-learning method was used to determine the control strategy of this system. They compared the operation of PV-priority and Q learning optimal control. The results illustrated that the adaptive controller held back the stored energy in order to be able to meet the critical loads in the future when the electricity price is higher, which can shave the peak electricity demand by about 50%.

Sukamongkol et al. [226] developed a simulation model for PV system under certain load requirements and meteorological conditions. This model contains of PV array, battery, controller, inverter and a load model. The PV array power output model is the same as the model in [221]. The weather condition data is sent into the PV array model to determine the power generated data. At the same time, the load model will determine the power demand. The controller model will regulate the operation mode of the PV panel to determine when to power the load and when to charge the battery. In additional, this controller will determine the operation of batteries.

5.1.4 Battery storage Modeling

For the purpose of peak demand shaving and dispatching, numerous studies on the electricity storage are being reported in many publications. The benefit of coupling battery system with on-

site generation has been shown in [227], where electricity generated in PV panel can either be used by building or stored in battery which enhances the economic potential of both systems beyond a simple sum of benefits that one might expect.

Nair and Garimella [228] discussed and assessed battery energy storage technologies from a technical and economic perspective in SIMULINK. An overview of various small-scale energy storage technologies followed by a modeling framework to assess the benefits of battery as an integral component of a renewable energy technology, such as PV systems, was provided in this paper. The battery system in this paper was modeled by a generic function block obtained from the *SimPowerSystems* toolbox in Simulink. Results from this simulation model showed that NiMH batteries have the highest potential for development in small-scale on-site energy integration applications. In spite of a poor technical performance, affordability and availability are two factors leading to dominant use of lead-acid batteries in renewable energy systems. Leadbetter and Swan [229] presented a battery energy system modeling approach for electricity demand peak shaving. Five-minute time step electricity profiles were fed into the energy storage model with the objective of reducing the peak electricity demand by optimization of the schedule of charging and discharging. They concluded that Peak demand reductions of between 42% and 49% were achieved by using 5 kWh or 8 kWh capacity battery energy storage systems. McKenna et al. [230] developed a novel battery model to study the economic impact of the battery for the occupants takes into account current U.K. feed-in tariff arrangements. The model can estimate the battery efficiency under varying rates of charge and discharge, as well as varying states of charge. They also quantified the operational energy losses by using the concept of voltage efficiency and coulombic. The overall energy efficiency of the battery can be viewed as the product of the battery's voltage and coulombic efficiencies.

A fundamental transport based model of lithium ion battery has been used to create a battery module based on Gao equivalent circuit model [231] in ESP-r for residential energy storage by Darcovich et al. [232]. This study evaluated the annual residential energy use in a typical Canadian home connected to the electrical grid, equipped with a micro-cogeneration system consisting of a sterling engine for supplying heat and power, coupled with a nominal 2 kW/6 kWh lithium ion battery. It was found that the complex nature of a multi-piece micro-cogeneration system benefits to a great extent from usage scenario simulations due to the specific capacities and outputs of the components. In the scenario tested in this study, the battery can reduce daily power consumption from electrical grid by 30%.

Based on these existing studies, the energy generation and storage modeling research is proposed in the following approach.

5.2 Other Building Energy Forecasting Model: RC Model

As introduced in the background section, resistance and capacitance (RC) network is a common simplified Physics based model for building energy consumption estimation. In order to compare the system identification model performance with different models, and increase the data fusion accuracy, RC model is proposed to be developed in this section.

Figure 5-1 shows the schematics of the RC model. Rs and Cs networks represent the building roof(s), external walls, and internal mass. External walls were modeled respectively according to their orientation: east wall, south wall, west wall, and north wall, due to the solar effect.

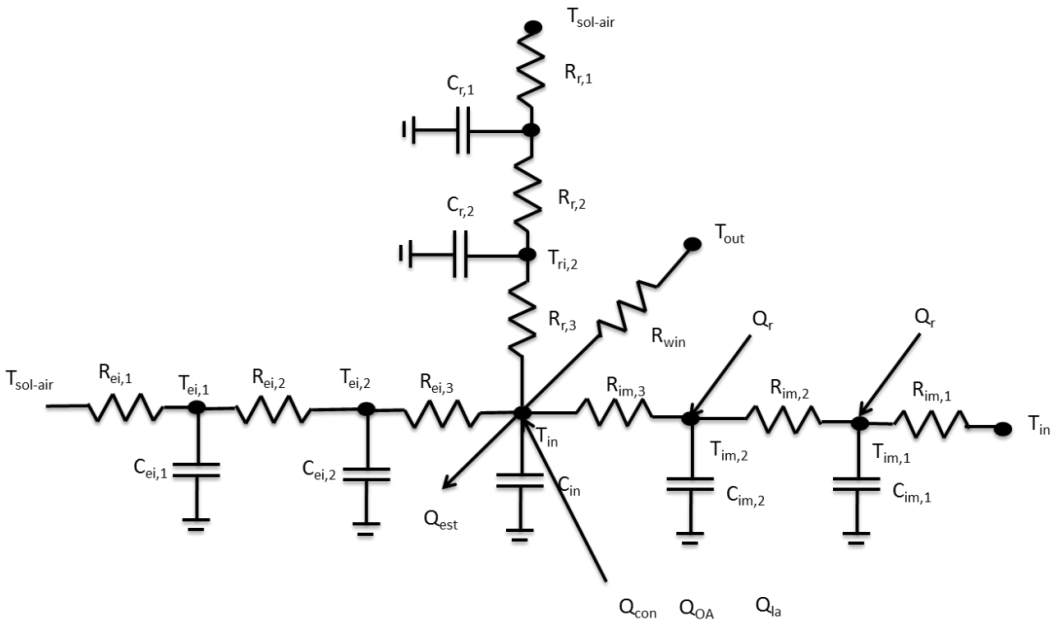


Figure 5-1. RC model structure for building energy estimation

At the first step, the RC model was developed just for the third floor for the simplification. All the external walls, internal masses, and roof are modeled as 3R2C structure. As shown in the schematics, all the Rs and Cs are parameters need to identify, while solar air temperature ($T_{sol\text{-}air}$), outdoor air temperature (T_{out}), solar transmitted heat gains (Q_{trans}), etc. are input (boundary) variables. Therefore the state variables and inputs variables are vectored in Eq. 5.19 and Eq. 5.20:

$$X = \left[T^{j}_{e,k,i} \quad T_{im,k,i} \right]^{T} \qquad \text{Eq. 5.19}$$

$$u = \left[T_{in,i} \quad T^{j}_{sol-air,i} \quad T_{out,i} \quad Q_{r1,i} \quad Q_{r2,i} \right]^{T} \qquad \text{Eq. 5.20}$$

Where,

j is the index for envelope including east, south, west, north, and roof, k is the index for surface sides including external side (as 1) and internal side (as 2), *i* is the index for time;

$T^{j}_{e,k,i}$ is the surface temperature of the external envelope for given surface *j*, and time *i*;

$T_{im,k,i}$ is the average surface temperature of internal mass for given time *i*;

$T_{in,i}$ is the average building indoor temperature for given time *i*;

$T^{j}_{sol-air,i}$ is the solar air temperature envelope *j*, and time *i*;

$T_{out,i}$ is the outside dry bulb air temperature for time *i*;

$T_{att,i}$ is the attatic temperature for time i, which is unconditioned;

$Q_{r,i}$ is the radiation heat gains to internal mass, which contains heat gains from solar remission and internal equipment radiation;

$Q_{con,i}$ is the convection heat gains to internal mass from indoor equipment;

$Q_{OA,i}$ is the heat gains from outdoor air, including infiltration and ventilation.

Apply the heat transfer theory on Figure 5-1, the energy balance equation of building can be described by:

$$C_{in}\frac{dT_{in}}{dt} = \sum_{j=1}^{5}\frac{T^{j}_{ei,2} - T_{in}}{R^{j}_{ei,3}} + \frac{T_{im,2} - T_{in}}{R_{im,3}} + \frac{T_{out} - T_{in}}{R_{win}} + Q_{con} + Q_{OA} \qquad \text{Eq. 5.21}$$

where, $C_{in}\frac{dT_{in}}{dt}$ is the indoor temperature changing rate multiple by the thermal capacitance, which equals to the heat gains/loos of the building; C_{in} is the capacitance associated with internal zone mass (air and furnishings); The first term at right hand side is the heat transfer rate from all the external envelops, including east, south, west, north walls and roof; the second term is the heat transfer rate from the internal envelopes; the third term is the heat transfer through the window, excluding the solar transmission; Q_{con} and Q_{OA} are the heat convection and outdoor air heat gains.

Besides this whole building heat balance equation, there is one heat balance equation at each node. For example, the heat balance for external envelopes can be modeled by Eq. 5.22 and Eq. 5.23:

$$C_{ei,1}^{j}\frac{dT_{ei,1}^{j}}{dt}=\frac{T_{sol-air}^{j}-T_{in}}{R_{ei,1}^{j}}-\frac{T_{ei,1}^{j}-T_{ei,2}^{j}}{R_{ei,2}^{j}} \quad \text{Eq. 5.22}$$

$$C_{ei,2}^{j}\frac{dT_{ei,2}^{j}}{dt}=\frac{T_{ei,1}^{j}-T_{ei,2}^{j}}{R_{ei,2}^{j}}-\frac{T_{ei,2}^{j}-T_{in}}{R_{ei,3}^{j}} \quad \text{Eq. 5.23}$$

Therefore, by rearranging the heat balance equations, a state space format model can be developed as Eq. 5.24 and Eq. 5.25:

$$\frac{dX}{dt}=AX+Bu \quad \text{Eq. 5.24}$$

$$Y=CX+Du \quad \text{Eq. 5.25}$$

Where, X and u are state vector and input vector, which as described before. A, B, C, and D are parameter matrices which are calculated by the Rs and Cs. In the case studied in this project, A is a 12×12 matrix, B is a 12×11 matrix, C is a 1×12 matrix, and D is a 1×11 matrix. The nonzero elements of these four coefficient matrices and vectors are:

$$A(1,1)=\frac{-1}{R_{ei,1}^{s}C_{ei,1}^{s}}+\frac{-1}{R_{ei,2}^{s}C_{ei,2}^{s}} \qquad A(1,2)=\frac{1}{R_{ei,2}^{s}C_{ei,1}^{s}}$$

$$A(2,1)=\frac{1}{R_{ei,2}^{s}C_{ei,2}^{s}} \qquad A(2,2)=\frac{-1}{R_{ei,2}^{s}C_{ei,2}^{s}}+\frac{-1}{R_{ei,3}^{s}C_{ei,2}^{s}}$$

$$A(3,3)=\frac{-1}{R_{ei,1}^{e}C_{ei,1}^{e}}+\frac{-1}{R_{ei,2}^{e}C_{ei,2}^{e}} \qquad A(3,4)=\frac{1}{R_{ei,2}^{e}C_{ei,1}^{e}} \quad \text{Eq. 5.26}$$

$$A(4,3)=\frac{1}{R_{ei,2}^{e}C_{ei,2}^{e}} \qquad A(4,4)=\frac{-1}{R_{ei,2}^{e}C_{ei,2}^{e}}+\frac{-1}{R_{ei,3}^{e}C_{ei,2}^{e}}$$

$$A(5,5) = \frac{-1}{R^n_{ei,1}C^n_{ei,1}} + \frac{-1}{R^n_{ei,2}C^n_{ei,2}} \qquad A(5,6) = \frac{1}{R^n_{ei,2}C^n_{ei,1}}$$

$$A(6,5) = \frac{1}{R^n_{ei,2}C^n_{ei,2}} \qquad A(6,6) = \frac{-1}{R^n_{ei,2}C^n_{ei,2}} + \frac{-1}{R^n_{ei,3}C^n_{ei,2}}$$

$$A(7,7) = \frac{-1}{R^w_{ei,1}C^w_{ei,1}} + \frac{-1}{R^w_{ei,2}C^w_{ei,2}} \qquad A(7,8) = \frac{1}{R^w_{ei,2}C^w_{ei,1}}$$

$$A(8,7) = \frac{1}{R^w_{ei,2}C^w_{ei,2}} \qquad A(8,8) = \frac{-1}{R^w_{ei,2}C^w_{ei,2}} + \frac{-1}{R^w_{ei,3}C^w_{ei,2}}$$

$$A(9,9) = \frac{-1}{R^r_{ei,1}C^r_{ei,1}} + \frac{-1}{R^r_{ei,2}C^r_{ei,2}} \qquad A(9,10) = \frac{1}{R^r_{ei,2}C^r_{ei,1}}$$

$$A(10,9) = \frac{1}{R^r_{ei,2}C^r_{ei,2}} \qquad A(10,10) = \frac{-1}{R^r_{ei,2}C^r_{ei,2}} + \frac{-1}{R^r_{ei,3}C^r_{ei,2}}$$

$$A(11,11) = \frac{-1}{R_{im,1}C_{im,1}} + \frac{-1}{R_{im,2}C_{im,2}} \qquad A(11,12) = \frac{1}{R_{im,2}C_{im,1}}$$

$$A(12,11) = \frac{1}{R_{im,2}C_{im,2}} \qquad A(12,12) = \frac{-1}{R_{im,2}C_{im,2}} + \frac{-1}{R_{im,3}C_{im,2}}$$

$$B(1,2) = \frac{1}{R^s_{ei,1}C^s_{ei,1}} \qquad B(2,1) = \frac{1}{R^s_{ei,3}C^s_{ei,2}}$$

$$B(3,3) = \frac{1}{R^e_{ei,1}C^e_{ei,1}} \qquad B(4,1) = \frac{1}{R^e_{ei,3}C^e_{ei,2}}$$

$$B(5,4) = \frac{1}{R^n_{ei,1}C^n_{ei,1}} \qquad B(6,1) = \frac{1}{R^n_{ei,3}C^n_{ei,2}} \qquad \text{Eq. 5.27}$$

$$B(7,5) = \frac{1}{R^w_{ei,1}C^w_{ei,1}} \qquad B(8,1) = \frac{1}{R^w_{ei,3}C^w_{ei,2}}$$

$$B(9,6) = \frac{1}{R^r_{ei,1}C^r_{ei,1}} \qquad B(10,1) = \frac{1}{R^r_{ei,3}C^r_{ei,2}}$$

$$B(11,1) = \frac{1}{R_{im,1}C_{im,1}} \qquad B(11,8) = \frac{1}{C_{im,1}}$$

$$B(12,1) = \frac{1}{R_{im,3}C_{im,2}} \qquad B(12,9) = \frac{1}{C_{im,2}}$$

$$C(1,2) = \frac{A_s}{R^s_{ei,3}} \qquad C(1,4) = \frac{A_e}{R^e_{ei,3}} \qquad C(1,6) = \frac{A_n}{R^n_{ei,3}}$$
$$C(1,8) = \frac{A_w}{R^w_{ei,3}} \qquad C(1,10) = \frac{A_r}{R^r_{ei,3}} \qquad C(1,12) = \frac{A_{im}}{R_{im,3}} \quad \text{Eq. 5.28}$$

$$D(1,1) = -(\frac{A_s}{R^s_{ei,3}} + \frac{A_e}{R^e_{ei,3}} + \frac{A_n}{R^n_{ei,3}} + \frac{A_w}{R^w_{ei,3}} + \frac{A_r}{R^r_{ei,3}} + \frac{A_{im}}{R_{im,3}} + \frac{A_{win}}{R_{win}})$$
$$D(1,7) = \frac{A_{win}}{R_{win}} \quad \text{Eq. 5.29}$$

Finally, the building heat gains can be estimated by:

$$Q_{gain} = Y + Q_{con} + Q_{OA} \quad \text{Eq. 5.30}$$

Therefore, substituting Eq. 5.30 into Eq. 5.31:

$$C_{in}\frac{dT_{in}}{dt} = Q_{gain} - Q_{HVAC} \quad \text{Eq. 5.31}$$

If the zone temperature is controlled constant when HVAC system provide enough cooling into the building, the sensible heating gains (Q_{gain}) is equal to the HVAC cooling providing (Q_{HVAC}). However, if the HVAV system is off or building heat gains exceeds HVAC system capacity, the room temperature will change. Therefore, two different operation scenarios have been considered in study:

1. zone temperature is maintained at a constant around the temperature setpoint;

2. zone temperature is floating with HVAC system is off or heat gains exceed HVAC system capacity.

In case one, the left hand side of Eq. 5.31 is zero, and Q_{gain} is equal to Q_{HVAC}. By solving Eq. 5.24, Eq. 5.25 and Eq. 5.30, Q_{HVAC} can be determined. In case two, Q_{HVAC} is known (0 or maximum capacity), then the zone temperature can be back calculated from solving Eq. 5.24, and Eq. 5.25. In this RC model, it is very importance to determine the coefficients (Rs and Cs). It is unfeasible to calculate each resistance and capacitance from the Physics theory. As a preliminary method testing in this step, just the first case has been considered, while the second case will be included in the future work. This test is based on the reference medium size office building EnergyPlus model provided by DOE. The detailed information has been introduced in section 3.5.1. The overall window area is 135 m^2, and the areas external envelopes are 309 m^2, 139 m^2, 309 m^2 139 m^2, and 618 m^2 for south wall, east wall, north wall, west wall, and roof, respectively. The internal mass surface are is 3321 m^2. The occupant density is assumed to be 18.5 m^2/person from 8 am to 6 pm. Equipment and light are set to be active one hour prior to occupancy. The lighting density is 16.1 W/m^2, and equipment density is 10.8 W/m^2. The internal heat gains are split in to convective part and radiative part. 40% of the light heats will go to return air, 40% is radiative heat gains and the remaining 20% is the convective heat gains. For the equipment heat gains, 50% is radiative and 50% is convective. The training period is Aug. 2 to Aug. 15, and the forecasting period is Aug 20 to Aug 22. The location of this building is Philadelphia, PA.

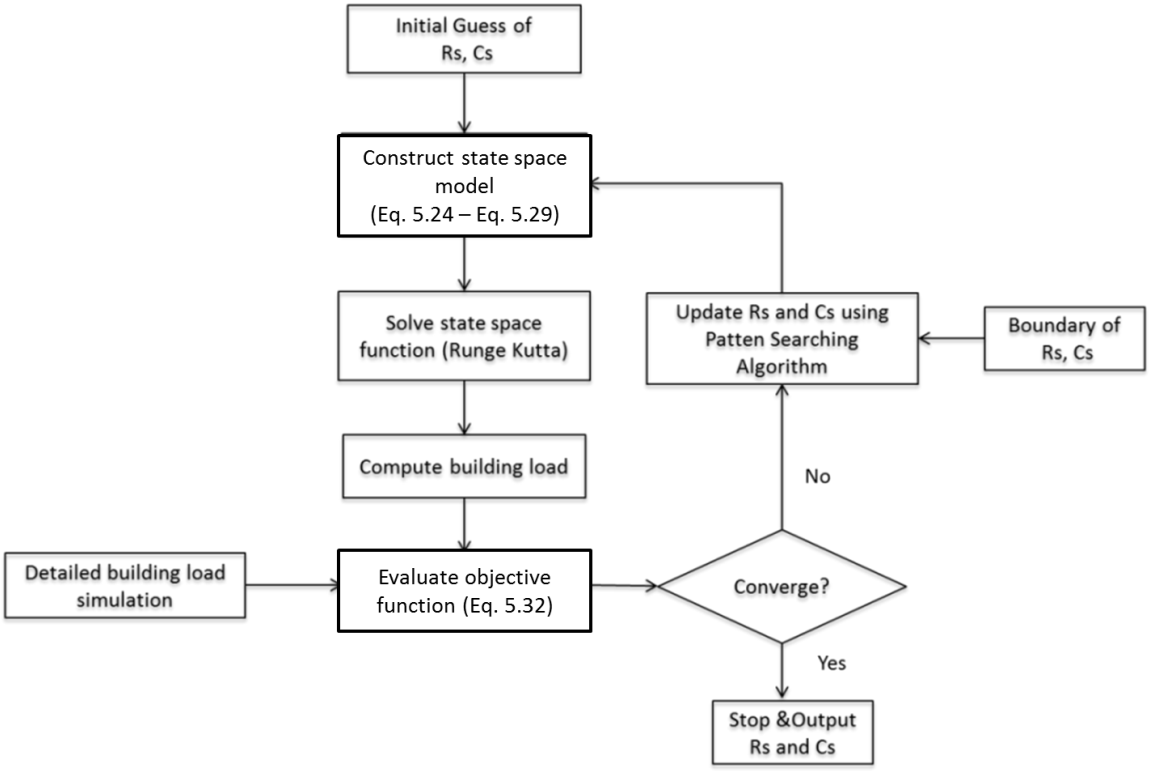


Figure 5-2. RC model training process flowchart

5.2.1 RC model Training

The model training process is to determine the Rs and Cs in the *A*, *B*, *C* and *D* matrices. They will be identified by comparing HVAC load estimated from simplified Physics model, RC model (Eq. 5.24 to Eq. 5.30), with the actual HVAC load from the detailed physics based simulation model (EnergyPlus model). The general procedure is shown in Figure 5-2. Optimization methods will be used to determining the parameters to minimize the difference between the HVAC load from RC model and EnergyPlus model. The objective function for this optimization problem is Eq. 5.32.

$$J_{RC} = \sqrt{\frac{\sum_{j=1}^{N}\left(Q_{RC,j} - Q_{Act,j}\right)^2}{N-1}} \qquad \text{Eq. 5.32}$$

Where, Q_{RC} and Q_{Act} are the HVAC loads (cooling and heating) from RC model and EnergyPlus model, respectively; N is the total time step of this whole simulation. Figure 5-2 shows the process for parameter identification. The initial guess and boundary of Rs and Cs are determined from the Physics theory. Runge Kutta method is used to solve the state space equations Eq. 5.24 and Eq. 5.25 to calculate the state variables, the estimated building load is then determined by Eq. 5.30. Pattern Searching optimization method is unitized here to update the Rs and Cs to minimize the objective function.

The performance if the simplified model is evaluated using normalized root-mean-square errors:

$$E_{rms} = \frac{100}{Q_{Act,m}} \sqrt{\frac{\sum_{j=1}^{N} \left(Q_{RC,j} - Q_{Act,j}\right)^2}{N-1}} \qquad \text{Eq. 5.33}$$

where, $Q_{Act,m}$ is the maximum load for the simulation period. The preliminary model training results are shown in Figure 5-3, and the overall normalized root-mean-square error is 11.3%. The normalized root-mean-square error for occupant period is 9.02%.

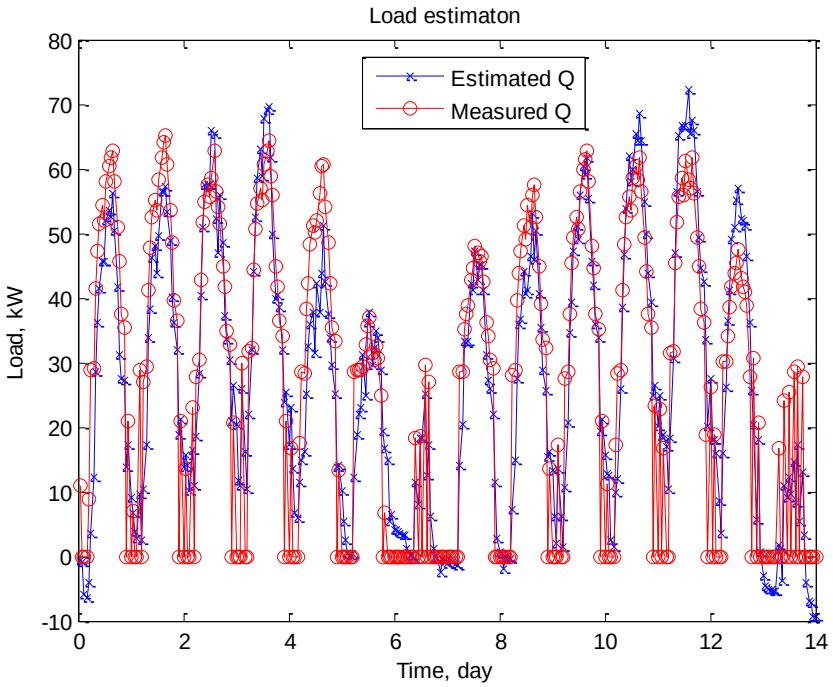


Figure 5-3. RC model training results

The estimation results from the RC model (blue line) capture the trend of the results from EnergyPlus (red line). The RC model tends to have larger errors when the load is high, while its overall accuracy is still acceptable. Different approaches will be adapted to improve the estimation accuracy. The future work of the RC model development will focus on improve its estimation accuracy and combining it with system identification energy estimation model.

5.2.2 RC model Forecasting

The RC model trained above is used to predict the building load of the same building under the same building operation scheme but different weather condition. The forecasting period is Aug 20 to Aug 22. Figure 5-4 shows the forecasting results. The measured results are from EnergyPlus simulation and estimated results area from the RC model. The overall rMSE is 13.17%. However, the load forecasting accuracy is much higher during day time when HVAC

system is on than that at night when system is off. Therefore further improvement will try to separate these two periods.

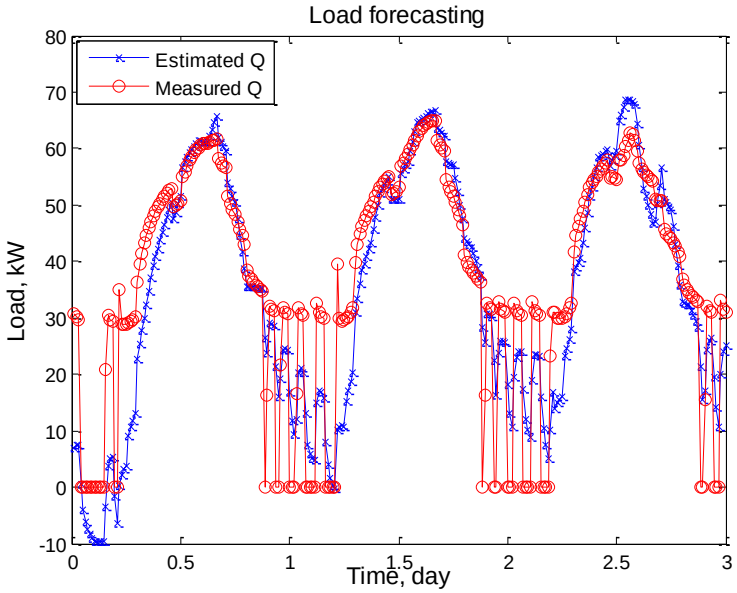


Figure 5-4. RC model forecasting results

RC model, as a grey box model, has been used in numerous building energy modeling studies. A preliminary study is conducted here to test the feasibility of RC model. The results from RC model prove it is a promising potential model structure for this study. In the future, the RC model will be served as a potential model structure in building energy forecasting and will also be integrated with system identification model through data fusion techniques to improve the overall energy forecasting accuracy and robustness.

5.3 Ice Tank Thermal Storage Model Development

In this section, the ice tank thermal storage model is developed based on the previous literature. The overall model structure is published and validated in [233]. The model will be introduced in the following sections, and the parameters in this model are identified based on the EnergyPlus simulation results.

As introduced in [233], the state of charge (SOC) of the ice tank is calculated as Eq. 5.34:

$$SOC_{t+\Delta t} = \begin{cases} SOC_t + \dfrac{u_t \Delta t}{Cap_s}, & charging \\ SOC_t - \dfrac{u_t \Delta t}{Cap_s}, & discharging \end{cases} \quad \text{Eq. 5.34}$$

where SOC_t is the state of charge at time t; u_t (Btu/h) is the charging/discharging rate at time t; Δt (h) is the simulation time interval, which is 1 h; Cap_s (Btu) is the capacity for the ice storage tank. The charging/discharging rate is determined by the heat transfer between the chilled water and the storage ice, which is dependent on the inlet fluid temperature (T_{chws}), secondary fluid flow rate (m_{chw}), and heat transfer effectiveness ($\varepsilon_{c,t}$, $\varepsilon_{d,t}$):

$$u_{c,t} = \varepsilon_{c,t} m_{chw} c_f (T_s - T_{chws}) \quad \text{Eq. 5.35}$$

$$u_{d,t} = \varepsilon_{d,t} m_{chw} c_f (T_{chws} - T_s) \quad \text{Eq. 5.36}$$

During the charging period, the inlet fluid temperature, T_{chws}, and flow rate, m_{chw} are determined by the dedicated chiller, while during the discharging period, they are determined by the building chilled water loop, such as the coils. Braun presented that heat transfer effectiveness ($\varepsilon_{c,t}$ and $\varepsilon_{d,t}$) is related with the ice tank state of charge, SOC:

$$\varepsilon_{c,t} = a + a_1 \times SOC_{r,t} + a_2 \times SOC_{r,t}^2 + a_3 \times SOC_{r,t}^3 + a_4 \times SOC_{r,t}^4 + a_5 \times SOC_{r,t}^5 \quad \text{Eq. 5.37}$$

$$\varepsilon_{d,t} = b + b_1 \times SOC_{r,t} + b_2 \times SOC_{r,t}^2 + b_3 \times SOC_{r,t}^3 + b_4 \times SOC_{r,t}^4 \quad \text{Eq. 5.38}$$

Where, $SOC_{r,t}$ is relative state of charge, a and b are heat transfer effectiveness coefficients which need to be determined by parameter estimation. Braun determined these coefficients by fitting the heat transfer effectiveness curves between modeled results and experiment measurements. Therefore, based on the pervious study, the ice tank model will be developed with chilled water flow rate ratio and inlet temperature as inputs and with charging, discharging rate, and SOC as outputs.

Based on the model equations, the overall procedure of ice tank model development and validation is shown in Figure 5-5. Same ice tank operation signals will be applied in EnergyPlus model and the simplified physics model. Then the ice tank charging, discharging rate, chilled water inlet water temperature, and flow rate will be collected as measurements to develop and validated the simplified model.

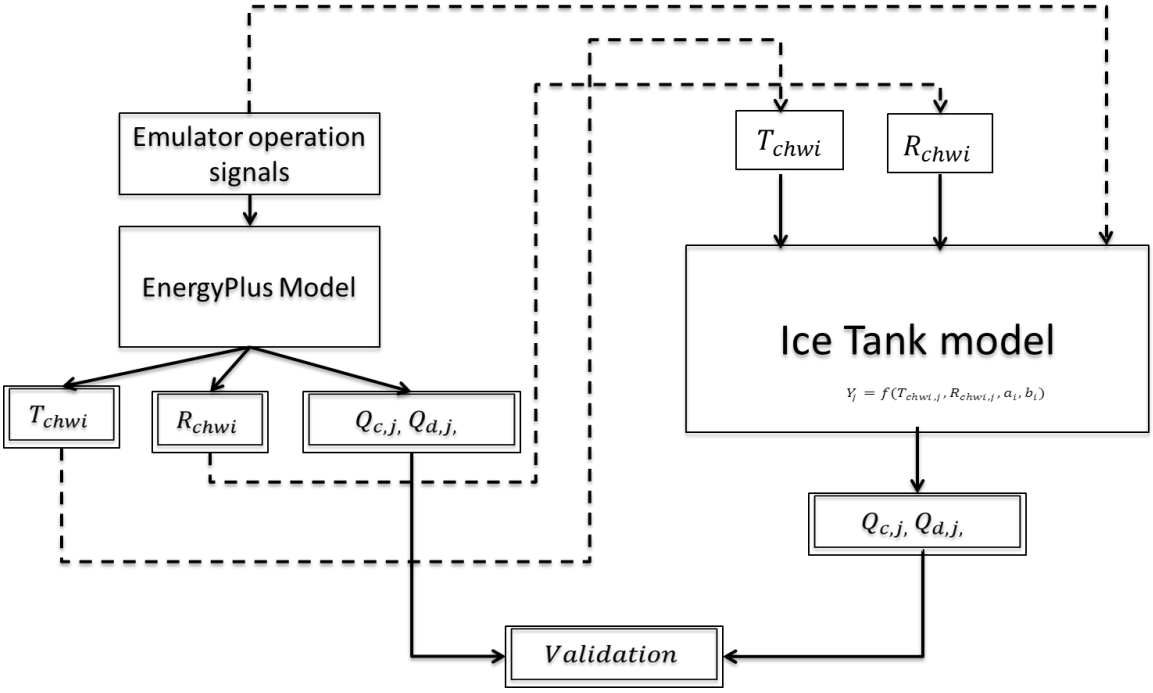


Figure 5-5. Ice tank model development and validation procedure

5.3.1 Ice Tank Thermal Energy Storage Model Development: model training

5.3.1.1 Model training method

As introduced above, the ice tank model will be developed follow the procedure presented by Braun, and the parameters in this model will first be identified based on the emulator simulation results.

Based on the discussion above, the overall ice tank performance function is expressed as:

$$Y_j = f_{ice}(T_{chwi,j}, R_{chwi,j}, a_i, b_i)$$

Where, Y_j is the output vector of the ice tank model, which contains charging rate, $Q_{c,}$, discharging rate, Q_d, and state of charge, SOC.

$$Y_j = [Q_{c,j,}Q_{d,j,}SOC_j]^T$$

R_{chw} is the ratio of the inlet chilled water flow rate to the maximum flow rate. a_i and b_i are the heat transfer effectiveness coefficients, which are identified through optimization approach. This optimization approach is to minimize the difference between the forecasted charging/discharging rate from the simplified model and those from EnergyPlus simulation:

$$J_{min}(a_i, b_i) = \sqrt{\frac{\sum_{j=1}^{N}(Q_{c,j} - \hat{Q}_{c,j})}{N-1}} + \sqrt{\frac{\sum_{j=1}^{N}(Q_{d,j} - \hat{Q}_{d,j})}{N-1}}$$

Where, *N* is the number of the total time step within whole training period, $\hat{Q}_{c,j}$ and $\hat{Q}_{d,j}$ are ice tank charging rate and discharging rate from EnergyPlus simulation, respectively. Pattern search optimization method is applied here to minimize the objective function by changing the parameters: a_i and b_i.

5.3.1.2 Model training settings

The ice tank training data here is generated in a validated EnergyPlus model for a 1-story building with an ice tank with capacity of 0.25 GJ. This EnergyPlus model is provided by US. DOE as a reference EnergyPlus model which has been validated against the real field measurement. In this chapter, this EnergyPlus model will be used in lieu as a real building. This building is divided into 4 exterior and 1 interior conditioned zones. The dedicated chiller is in series configuration, whose nominal capacity is 15 kW, and nominal COP is 3.2. The chilled water in the charging loop is 40% mixture of ethylene glycol and water. The specific heat of the chilled water is different at different temperature. However, to keep the simplicity of the on-line

model, just a constant specific heat is used, which is 3.38 $kJ/(K \cdot kg)$ for the charging loop and 3.44 $kJ/(K \cdot kg)$ for the discharging loop. These two numbers are the specific heat of the chilled water at its setpoints in charging and discharging process, -5 °C and 7 °C. Since the chilled water temperature is controlled by the flow valve, it will not be far away from the setpoints.

Table 5-1. Ice tank operation scheme

Time	Operation
0 – 9 am	Ice tank is charged by dedicated chiller until its full
9am -10 pm	Ice tank is discharged to provide cooling to the building until its deplete
10pm -12 am	Ice tank is charged by dedicated chiller

As introduced in Chapter 2, the operation of the ice tank (charging, discharging and dominant) is controlled by the temperature setpoint of the dedicated chiller. The training period is August 1st and August 2nd, and the location of this building is Philadelphia. The general operation schedule is summarized in Table 5-1.

Model training results

Based on the training settings, the ice tank was controlled to charge and discharge within two days. The optimization results for the heat transfer coefficient identification are shown in Table 5-2.

Table 5-2. Heat transfer effectiveness coefficient identification results

Coefficient	Value	Coefficient	Value
a_0	0.89	b_1	0.11
a_1	-0.61	b_2	0.82
a_2	0.86	b_3	0.83
a_3	-0.93	b_4	-0.75
a_4	0.92	b_5	-0.43

a_5	-0.3	b_6	0.24
b_0	0.22	b_7	0.02

Therefore the heat transfer effectiveness function can be illustrated as:

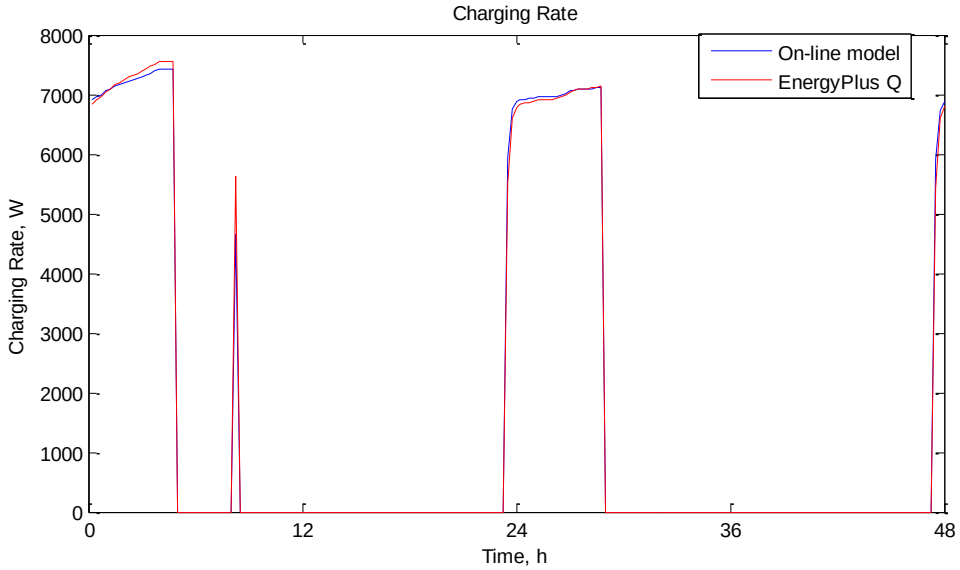
$$\varepsilon_{c,t} = 0.89 + 0.61 \times SOC_{r,t} + 0.86 \times SOC_{r,t}^2 - 0.93 \times SOC_{r,t}^3 + 0.92 \times SOC_{r,t}^4 - 0.3 \times SOC_{r,t}^5 \quad \text{Eq. 5.39}$$

$$\varepsilon_{d,t} = 0.22 + 0.11 \times SOC_{r,t} + 0.82 \times SOC_{r,t}^2 + 0.83 \times SOC_{r,t}^3 - 0.75 \times SOC_{r,t}^4 - 0.43 \times SOC_{r,t}^5 + 0.24 \times SOC_{r,t}^6 \quad \text{Eq. 5.40}$$

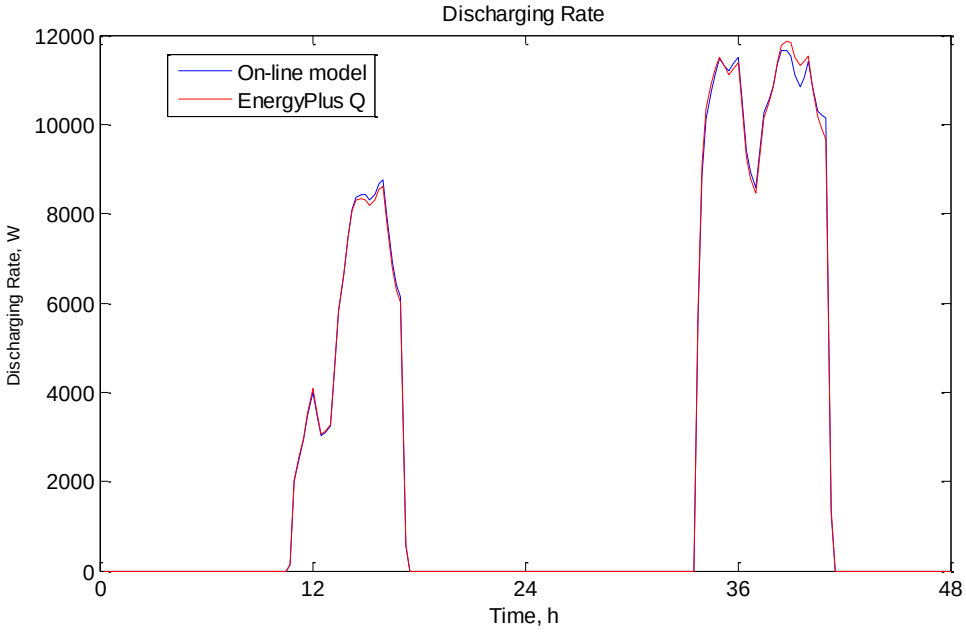
The overall ice tank operation training results are shown in Figure 5-6. The accuracy of the training process is shown in Table 5-3. In the model training process only the differences of charging and discharging rate are in the objective function to be minimized. The error of charging rate (Figure 5-6a) and discharging rate (Figure 5-6b) are very small, while that of SOC (Figure 5-6c) is relative large. Fortunately, when this ice tank connects with the building models, only the charging and discharging rate will affect the building operation, and the SOC is an internal variable which has no effect to the overall building cluster models.

Table 5-3. Ice tank model training and forecasting results

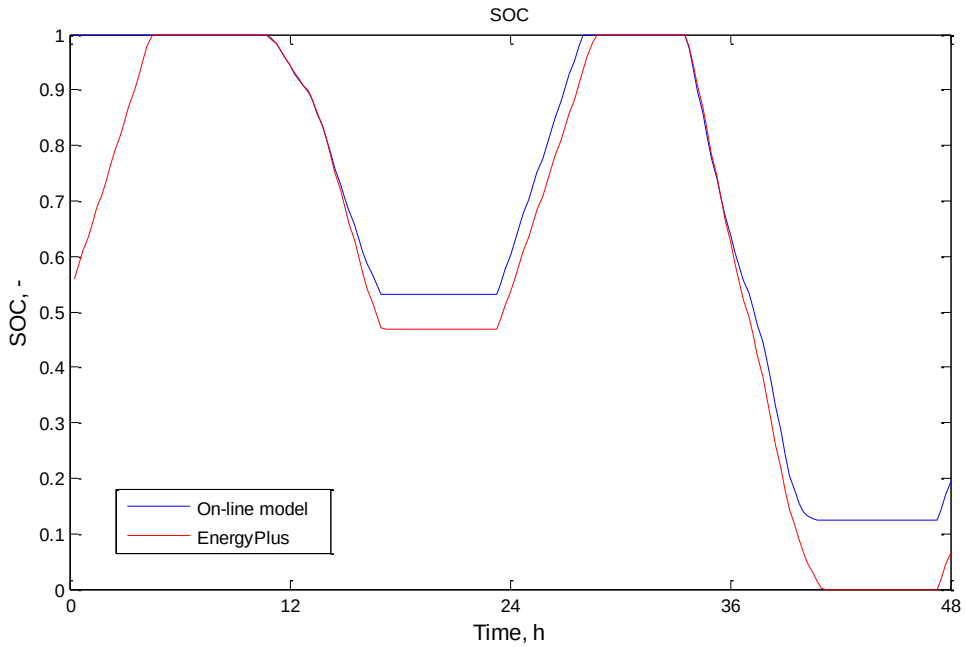
	Training Period		**Forecasting Period**	
Estimation	Accuracy, R^2	Figure	Accuracy, R^2	Figure
Chagrining rate	0.99	Figure 5-6a	0.98	Figure 5-7a
Discharging rate	0.96	Figure 5-6b	0.96	Figure 5-7b
State of Charge	0.89	Figure 5-6c	0.82	Figure 5-7c



a) Charging rate estimation



b) Discharging rate estimation



c) State of charge estimation

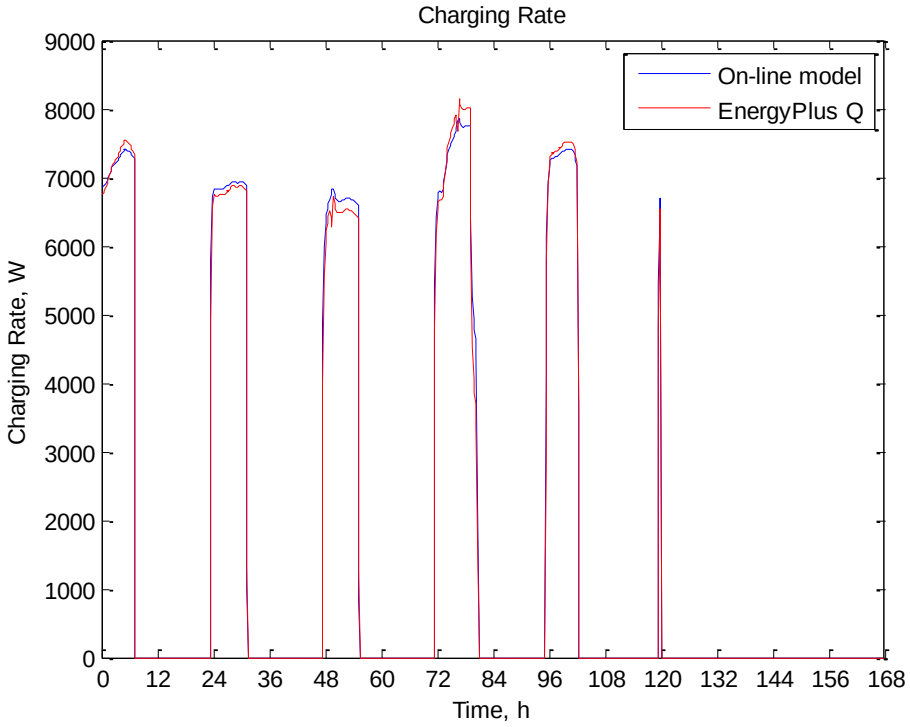
Figure 5-6. Ice tank model training results.

5.3.2 Ice Tank Thermal Energy Storage Model Development: model validation

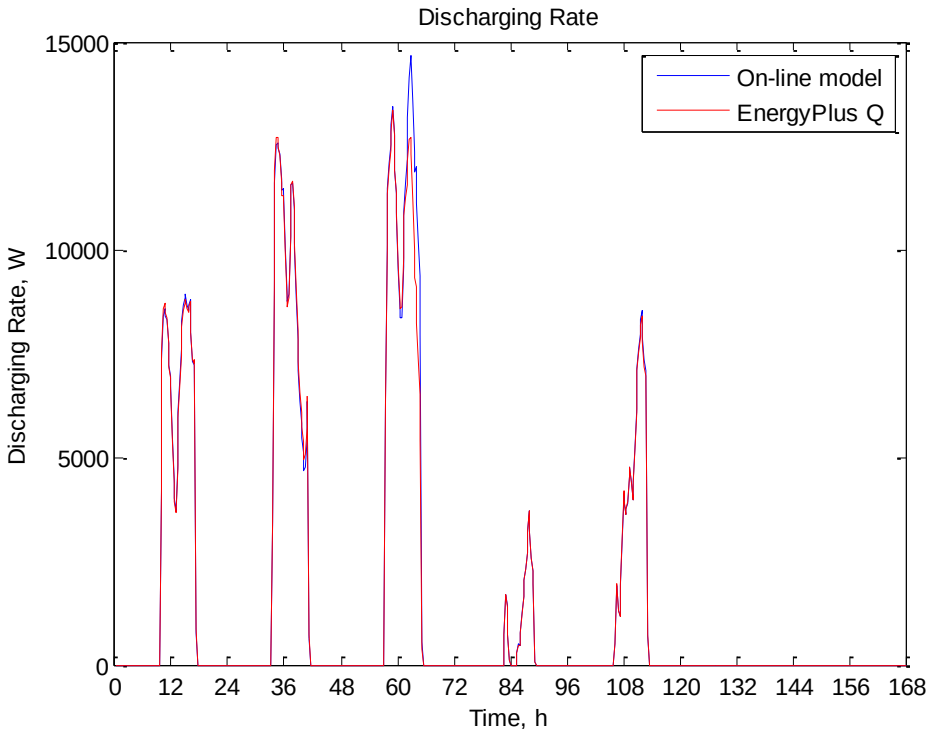
The ice tank thermal storage model developed in pervious sections will also be validated based on the same validated EnergyPlus simulation results for the following 7 days, from August 3rd to 9th. In the validation process, chilled water inlet temperature and flow rate ratio are the inputs of the ice tank model. The model outputs: charging rate, discharging rate will be validated against the EnergyPlus simulation results. The model settings during the model validation and forecasting period are exactly the same.

The forecasting results are shown in Figure 5-7, and the accuracy is also summarized in Table 5-3. The overall accuracy for the charging and discharging rate forecasting accuracy is above 95%. In Figure 5-7a, the charging rate forecasting from the on-line model is very close to the EnergyPlus simulation, while in Figure 5-7b, the discharging rate at early afternoon of the third day from the on-line model is higher than the EnergyPlus simulation results. The reason of this

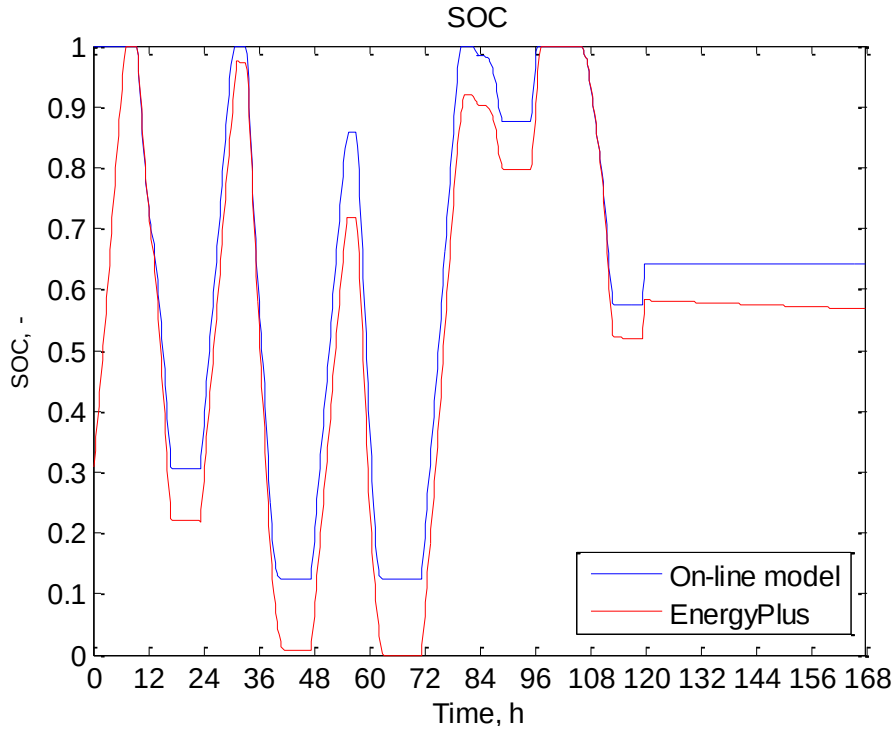
different is indicated in Figure 5-7c, where the SOC of the ice tank at the third day from the on-line model is much higher than that from EnergyPlus. There might be two potential reasons for this situation: 1) the tank loss is neglected in the current version of the on-line model, which will be added in next step; 2) the chilled water specific heat in the on-line model is higher than the real case. Hence a specific heat estimation model from temperature will also be developed in the following work. However, even with a constant specific heat, the model has already achieved above 95% accuracy.



a) Charging rate estimation



b) Discharging rate estimation



c) State of charge estimation

Figure 5-7 Ice tank model validation results.

5.4 PV Panel Model Development

A simplified PV panel model is developed and validated in this section based on previous studies. Similar to the ice tank model, this PV panel model will be firstly trained based on the emulator simulation data and then refined based on the real field experiment data. From previous study [130], the power output of the PV panel can be estimated as:

$$P_{pv} = [-(aG + b)(T_{db} + 0.03375G) + cG + d]M_s M_P \quad \text{Eq. 5.41}$$

Where G (W/m^2) is the absorbed solar irradiance on the PV panel, T_{db} (°F) is the dry-bulb ambient temperature; M_s and M_P are the number of modules in series and parallel, respectively. The coefficients *a*, *b*, and *c* will be identified based on EnergyPlus simulation results and real experiment measurements. The absorbed solar irradiance on the PV-panel can be estimated by Eq. 5.42:

$$G = G_{bt}(1 - F_B) + G_{dt}(1 - F_D) + G_r(1 - F_R) \quad \text{Eq. 5.42}$$

Where G_{bt}, G_{dt}, and G_r (W/m^2) are beam, diffuse and reflected hourly solar irradiance on the PV-panel, respectively, which are available at national solar radiation data [234] ; F_B, F_D, and F_R are angular losses factor for beam, diffuse and reflected irradiance respectively, which can be calculated by Eq. 5.43, Eq. 5.44 and Eq. 5.45 [235]:

$$F_B = \frac{exp\left(-\frac{cos\theta}{a_r}\right) - exp\left(-\frac{1}{a_r}\right)}{1 - exp\left(-\frac{1}{a_r}\right)} \quad \text{Eq. 5.43}$$

$$F_D = exp\left[-\frac{1}{a_r}\left[c_1\left(sin\beta + \frac{\pi - \frac{\beta\pi}{180} - sin\beta}{1 + cos\beta}\right) + c_2\left(sin\beta + \frac{\pi - \frac{\beta\pi}{180} - sin\beta}{1 + cos\beta}\right)^2\right]\right] \quad \text{Eq. 5.44}$$

$$F_R = exp\left[-\frac{1}{a_r}\left[c_1\left(sin\beta + \frac{\frac{\beta\pi}{180} - sin\beta}{1 - cos\beta}\right) + c_2\left(sin\beta + \frac{\frac{\beta\pi}{180} - sin\beta}{1 - cos\beta}\right)^2\right]\right] \quad \text{Eq. 5.45}$$

where, $c_1 = \frac{4}{3\pi}$; θ is the angle of solar incidence; β is the slope angle of inclined surface. a_r, *and* c_2 will also be estimated based against EnergyPlus simulation results and then real experiment measurements. Similar to the model development procedure for ice tank model, the emulator simulation results are used as measurements for the PV panel model developing and validation.

5.4.1 PV Panel Power Generation Model Development: model training

5.4.1.1 Model training method

The model structure introduced above has been developed for the PV power generation estimation. The overall PV panel power generation model can be expressed as Eq. 5.46:

$$Y_{PV} = f_{pv}(date, Ms, Mp, A, a, b, c, d, ar, c_2) \quad \text{Eq. 5.46}$$

Detailed Physics model in TRNSYS within the building cluster emulator developed in Chapter 2 is chosen to provide training and validation data. The parameters in Eq. 5.46, i.e. $a, b, c, d, ar,$ and c_2, will be identified using optimization approach , by minimizing the power generation estimation differences between the simplified model and detailed TRNSYS model. The objective function of this optimization problem is:

$$J_{min}(a,b,c,d,ar,c_2) = \sqrt{\frac{\sum_{j=1}^{N}(P_{pv,j}-\hat{P}_{pv,j})}{N-1}} \quad \text{Eq. 5.47}$$

Where, N is the number of the total time step within whole training period, $P_{pv,j}$ and $\hat{P}_{pv,j}$ are power generation estimation from simplified model and TRNSYS model, respectively. Pattern search optimization method is applied here to minimize the objective function by updating the parameters until the global minimize is achieved.

5.4.1.2 Model training settings

Sandia PV module model is used in TRNSYS model. There are a large number of empirical coefficients required to use the Sandia model. These coefficients are obtained after extensive measurements and data reduction. Sandia publishes a database for module and array performance parameters. The active area of one PV module is 0.63 m^2 (A in Eq. 5.46). There are 6 modules in series (M_S in Eq. 5.46), and 3 series strings in parallel (M_p in Eq. 5.46). The weather file used for the training data simulation is TMY file for Philadelphia, PA, and the training period is from June 1st to 14th. The hourly global solar radiation on the horizontal surface G_{th} and solar angles are provided by national solar radiation data [234]. As the first preliminary model training and test, the PV panel is placed horizontally. Therefore the solar radiation on the PV panel surface is identical to the solar radiation on horizontal surface. In next step, tilted (45°), facing south PV panel placement will be simulated.

5.4.1.3 Model training results

Based on the training settings, training data is generated for 14 days from TRNSYS model. The parameters in the simplified model have been identified based on the training data.

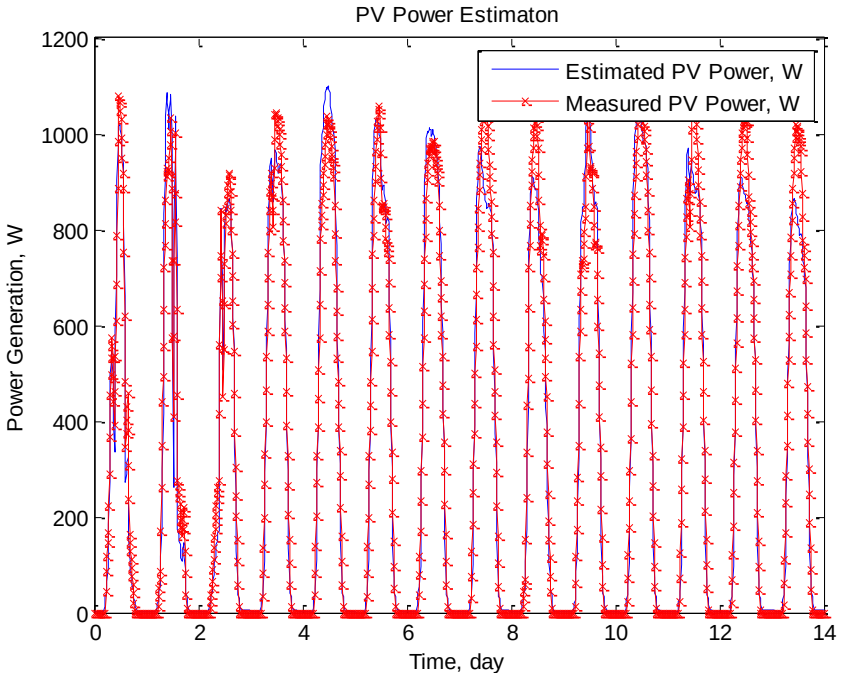
The performance if the simplified model is evaluated using normalized root-mean-square errors:

$$E_{rms} = \frac{100}{P_{Act,m}} \sqrt{\frac{\sum_{j=1}^{N} \left(P_{pv,j} - \hat{P}_{pv,j}\right)^2}{N-1}} \quad \text{Eq. 5.48}$$

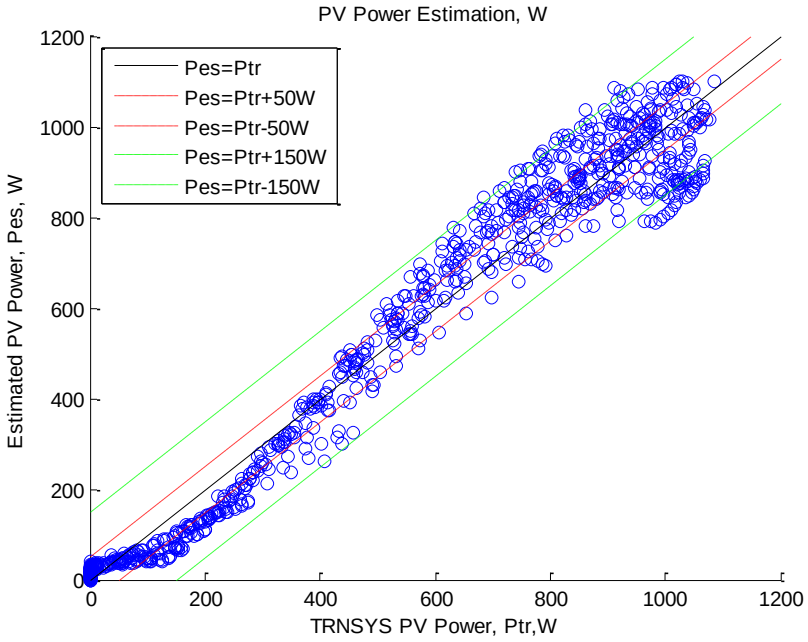
where, $P_{Act,m}$ is the maximum power generation simulation period, W. The parameter identification results are shown in Table 5-4. The overall PV panel power generation training results based on the parameters identified above are shown in Figure 5-8, and the overall normalized root-mean-square error is 4.61%, which is acceptable for the building power generation simulation requirement in this study.

Table 5-4. PV power generation model parameters identification results

Coefficient	Value	Coefficient	Value
a	0.0002	d	-0.6206
b	-0.0021	a_r	12.3450
c	0.0715	c_2	1.4939



a. power generation estimation



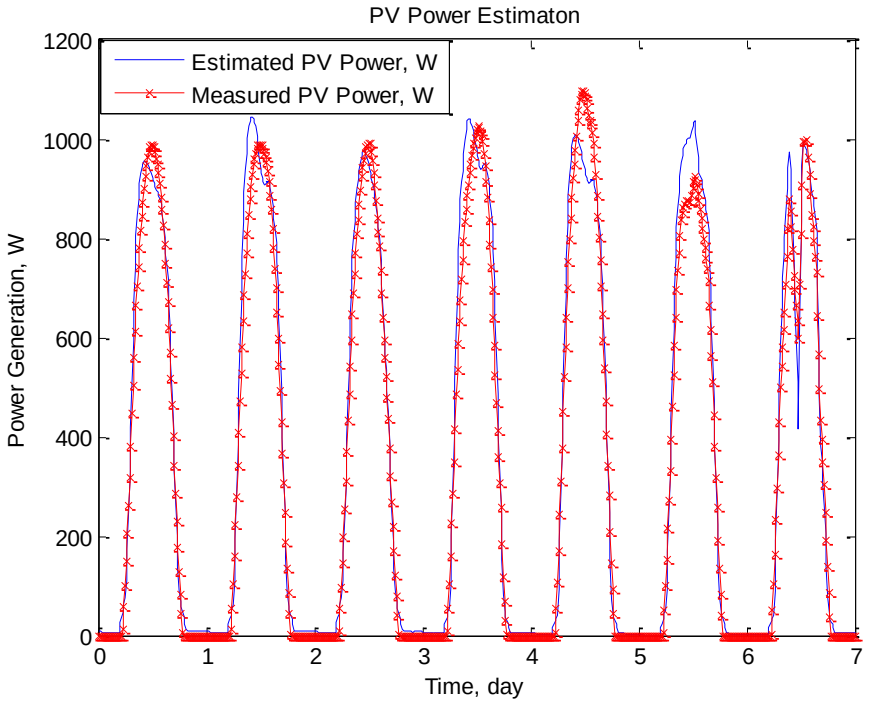
b. estimation accuracy

Figure 5-8. PV panel power generation training results

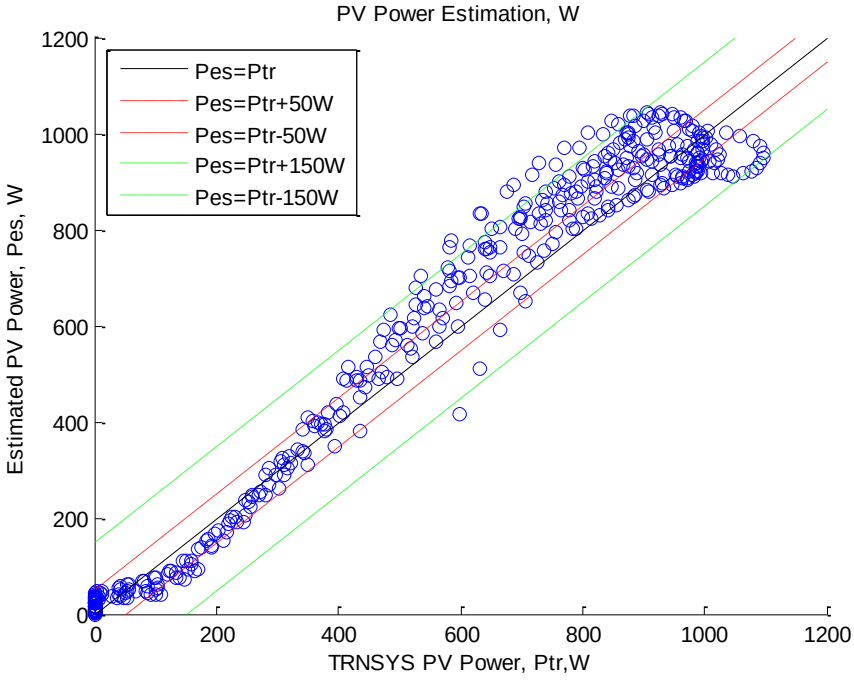
5.4.2 PV Panel Power Generation Model Development: model validation

5.4.2.1 Model validation method and results

The PV power generation model developed above will be validated and refined in this section based on the new validation data from the same TRNSYS model for the following 7 days simulation. In validation process, solar radiance and angle are also obtained from national solar radiation data. The power generation forecasting results are plotted in Figure 5-9.



a. Power generation forecasting



b. Forecasting accuracy

Figure 5-9. PV panel power generation forecasting results:

The overall normalized root-mean-square error is 5.85%. As shown in Figure 5-9b, the simplified PV power generation model has higher accuracy when power generation is small, while has relative lower accuracy when solar radiation is very strong. The reason for this lower accuracy is the cell temperature estimation error, because cell temperature is a factor with great influence on the PV panel power generation. In this simplified model a linear relationship is used to estimate the solar cell temperature from outdoor dry blub temperature and total solar radiation on the PV panel surface (Eq. 5.49).

$$T_{cell} = T_{db} + 0.03375G \qquad \text{Eq. 5.49}$$

This linear relationship has relative lower accuracy when solar radiation is strong, as shown in Figure 5-10. Future work will address the cell temperature estimation accuracy improvement, by updating the coefficient, or employing nonlinear equation.

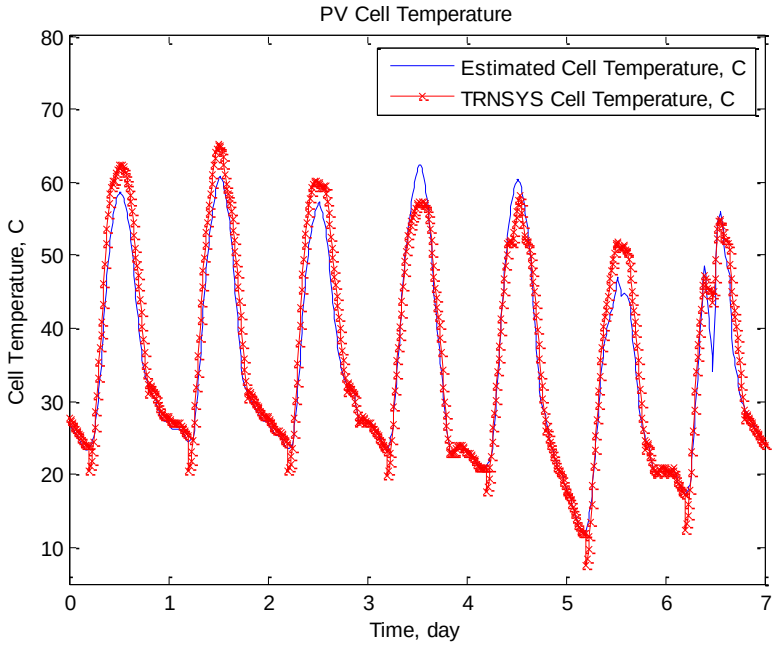


Figure 5-10. PV panel cell temperature forecasting results

5.5 Battery Model Development

As introduced in section 5.1.4, a simplified battery storage model will be developed based on [236]. The relationship between battery state of charge and charging/discharging current can be described as Eq. 5.50 and Eq. 5.51:

$$SOC_{t+\Delta t} = \begin{cases} SOC_t + \dfrac{I_{bat,t}\Delta t\,\eta_t}{C_{bat}}, & charging \\ SOC_t - \dfrac{I_{bat,t}\Delta t}{C_{bat}}, & discharging \end{cases} \quad \text{Eq. 5.50}$$

$$I_{bat,t} = \begin{Bmatrix} min\{\alpha C_{bat},\ (SOC_{max} - SOC_t) \times C_{bat}/(\eta_t \Delta t)\},\ charging \\ min\{\beta C_{bat},\ (SOC_{max} - SOC_t) \times C_{bat}/\Delta t\},\ discharging \end{Bmatrix} \quad \text{Eq. 5.51}$$

where SOC_t is the state of charge at time *t*; $I_{bat,t}$ (A) is the charging/discharging current at time *t*; Δt (h) is the simulation time interval, which is 1 h; C_{bat} (Ah) is nominal battery capacity; η_t is the round efficiency of the battery at time *t*. SOC_{max} is the maximum state of charge which is 1.0 in this research; SOC_{min} is the minimum state of charge which is 0.5 in this research.

The terminal voltage $V_{bat,t}$ (V) of the battery at time *t* is expressed in terms of its open circuit voltage, $V_{oc,t}$, and the voltage drop across the internal resistance of the battery (Eq. 5.52). $V_{oc,t}$ is calculated as Eq. 5.53:

$$V_{bat,t} = V_{oc,t} - I_{bat,t}R_{bat,t} \quad \text{Eq. 5.52}$$

$$V_{oc,t} = VF + b \times log(SOC_t) \quad \text{Eq. 5.53}$$

where *VF* (V) is the full charge rest voltage, *b* is an empirical constant, and $R_{bat,t}$ is the internal resistance of the battery, which is computed as Eq. 5.54:

$$R_{bat,t} = r_1 + r_2 \times SOC_t + \frac{1}{r_3 - r_4 \times SOC_t} \quad \text{Eq. 5.54}$$

where r_1, r_2, r_3, and r_4 are resistance coefficients.

The charging/discharging power for the battery $P_{bat,t}$ (W) are calculate as

$$P_{bat,t} = V_{bat,t} I_{bat,t} M_s M_p \quad \text{Eq. 5.55}$$

where M_s is number of battery unit in series, M_p is number of battery unit in parallel.

Above all, VF is a constant parameter provided by the battery catalog, which is 12.6 V at charging model or 12.4 V at discharging model. All the coefficients, such as b, r_1, r_2, r_3, and r_4 will be identified using emulation simulation results and then real experiment measurements

5.5.1 Battery Model Development: model training

Kinetic battery model in TRNSYS is used to simulate a battery to provide the on-line battery model training and validation data. This battery model is connected with the PV panel model and building model in the building cluster emulator. There are 10 modules in series (M_S in Eq. 5.55), and 10 series strings in parallel (M_p in Eq. 5.55). The capacity of the battery module is 86.1 Ah. Similar to the ice tank model training process, the overall battery performance function is expressed as:

$$Y_j = f_{BS}(I_{BS,j}, S_{BS,j}, r_{i,j}, b_j)$$

Where, Y_j is the output vector of the ice tank model, which contains charging rate, W, discharging rate, W, and state of charge.

$$Y_j = [P_{c,j,} P_{d,j,} SOC_{BS,j}]^T$$

$I_{BS,j}$ is the charging or discharging current. $S_{BS,j}$ is the battery operation state, 1: discharging, 2: changing, 0: dormant. $r_{i,j}$ and b_i are the battery parameters, which are identified through optimization approach. This optimization approach is to minimize the difference between the forecasted charging/discharging rate from the simplified model and those from EnergyPlus simulation:

$$J_{min}(r_i, b_i) = \sqrt{\frac{\sum_{j=1}^{N}(P_{c,j} - \hat{P}_{c,j})}{N-1}} + \sqrt{\frac{\sum_{j=1}^{N}(P_{d,j} - \hat{P}_{d,j})}{N-1}}$$

Where, N is the number of the total time step within whole training period, $\hat{P}_{c,j}$ and $\hat{P}_{d,j}$ are battery charging rate and discharging rate from battery model simulation, respectively. Pattern search optimization method is applied here to minimize the objective function by changing the parameters.

Table 5-5. Battery model parameter identification results

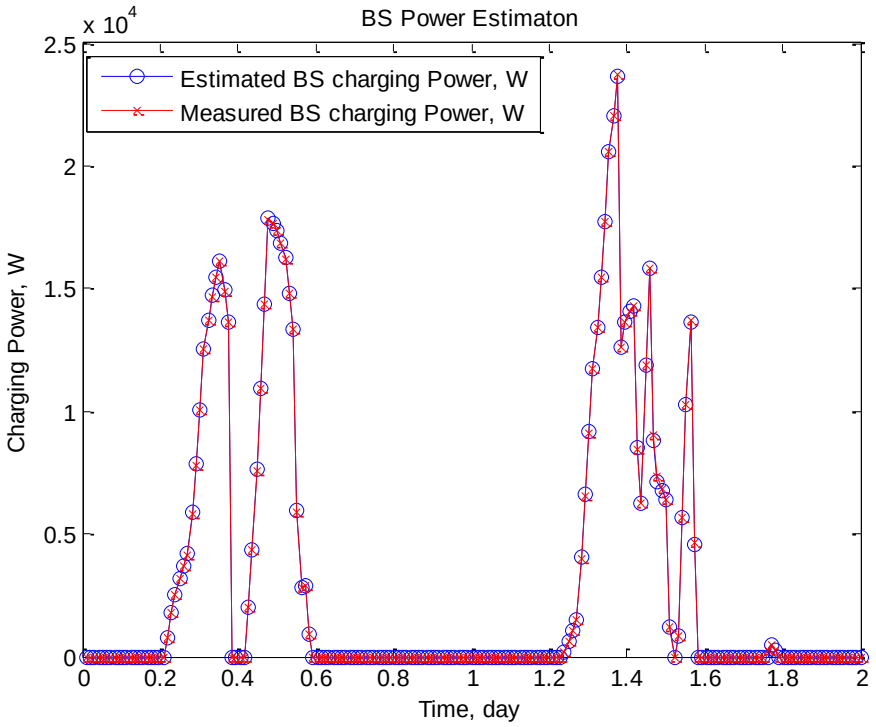
Coefficient	Value	Coefficient	Value
b_c	0.89	b_d	0.11
r_{1c}	-0.61	r_{1d}	0.82
r_{2c}	0.86	r_{2d}	0.83
r_{3c}	-0.93	r_{3d}	-0.75
r_{4c}	0.92	r_{4d}	-0.43

Based on the training settings, training data is generated for 4 days from TRNSYS model. The parameters in the simplified model have been identified based on the training data. Table 5-5 summarizes the parameters identification results. Using all these parameter the model training results are illustrated in Figure 5-11.

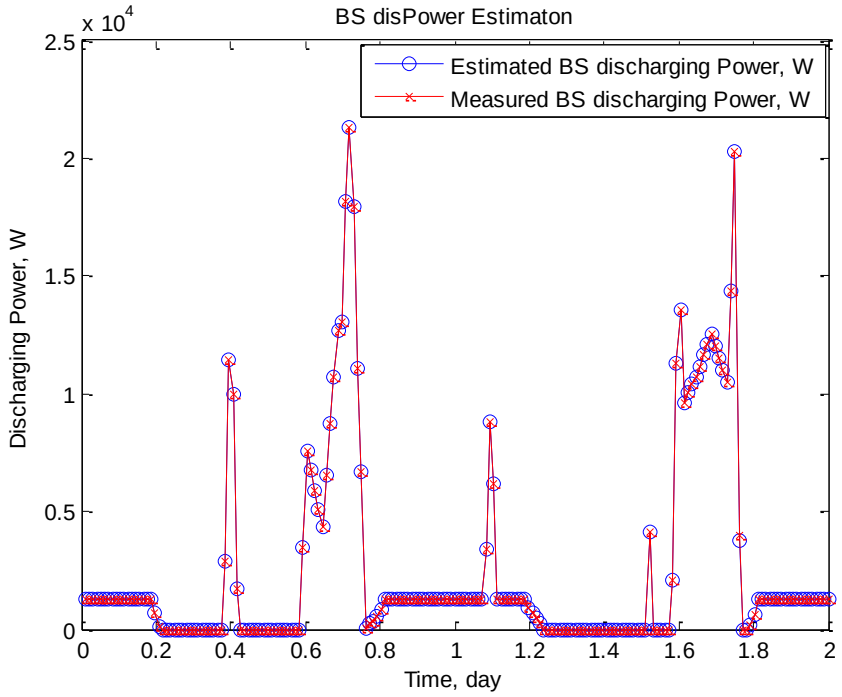
The performance of the simplified model is evaluated using normalized root-mean-square errors:

$$E_{rms} = \frac{100}{P_{Act,m}} \sqrt{\frac{\sum_{j=1}^{N}(P_{BS,j} - \hat{P}_{BS,j})^2}{N-1}} \quad \text{Eq. 5.56}$$

The overall model accuracy during the training period (E_{rms}) is 0.13% for the charging power, and is 0.07% for the discharging power. The state of charge is an internal variable and not in the objective function. Therefore the accuracy of state of charge is not calculated.

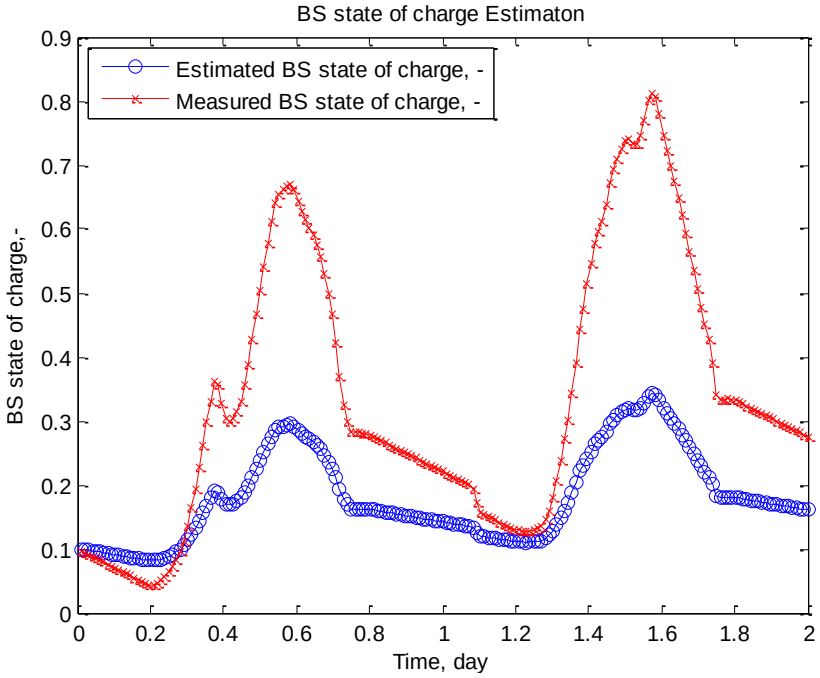


a. Charging rate estimation



b. Discharging rate estimation

c.



d. State of charge estimation

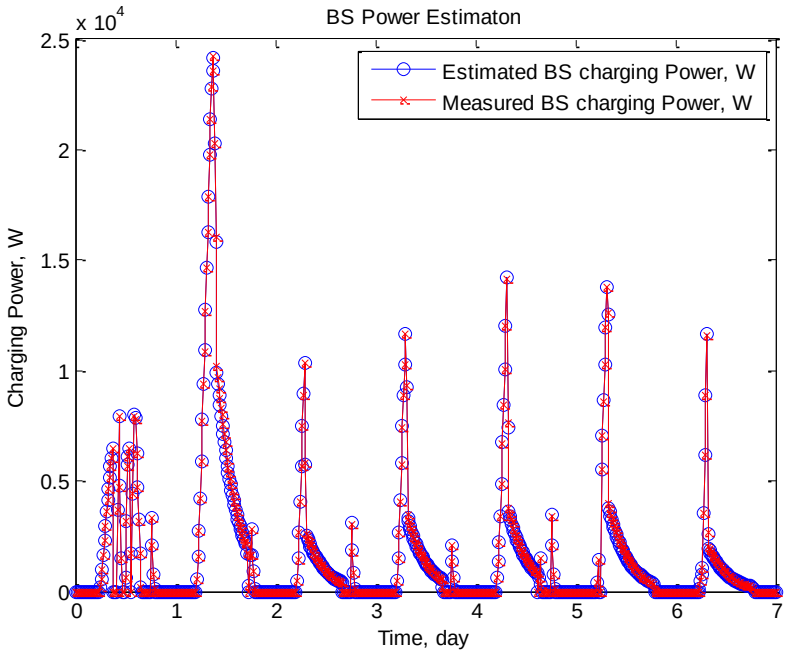
Figure 5-11 Battery model training results.

5.5.2 Battery Model Development: model validation

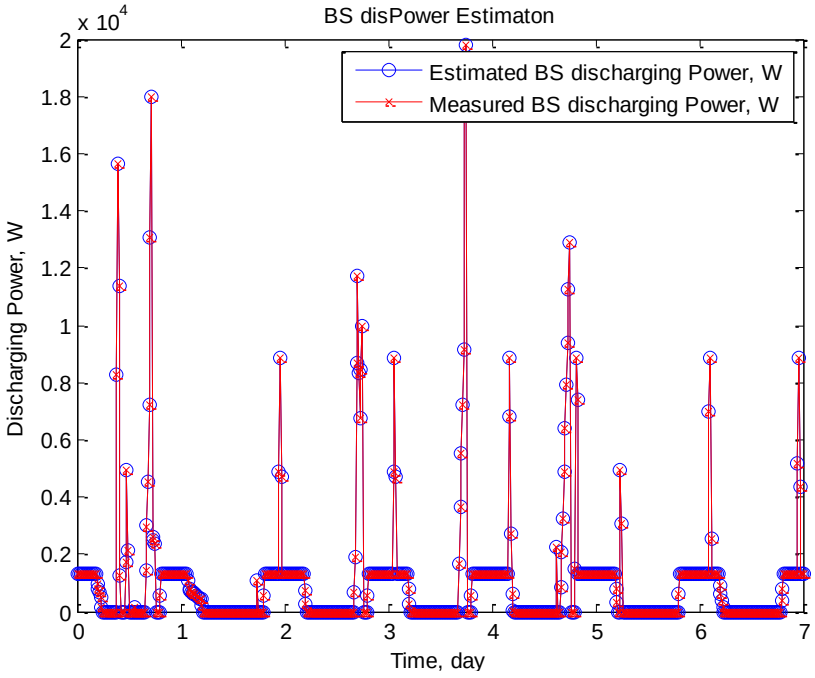
The battery electricity storage model developed in pervious section will also be validated based on the same validated TRNSYS simulation results for the following 7 days, from August 3rd to 9th. In the validation process, battery operation (charging/discharging) state and battery (charging/discharging) current are the inputs of the battery model. The model outputs: charging rate, discharging rate will be validated against the TRNSYS simulation results. The model settings during the model validation and forecasting period are exactly the same.

5.5.2.1 Model validation results

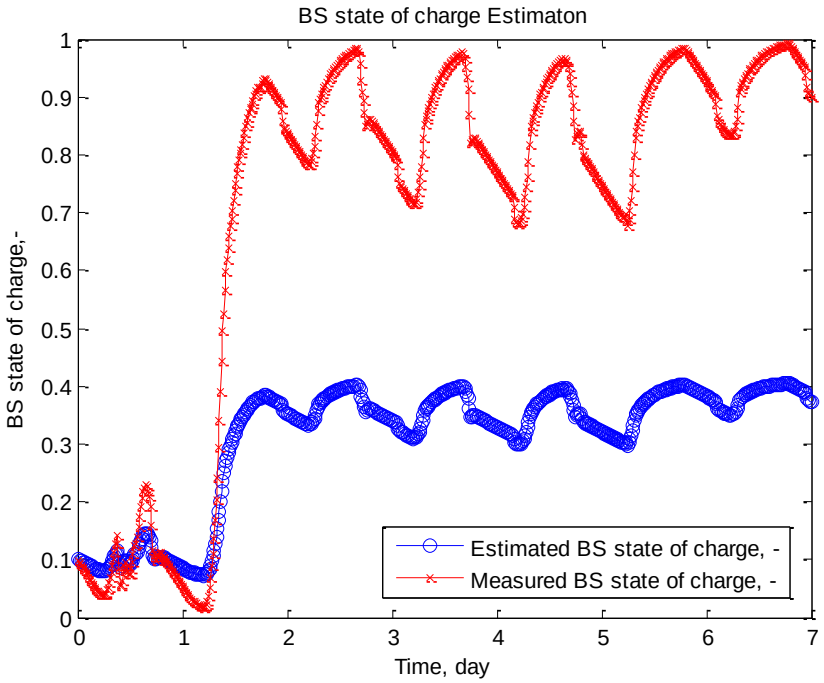
The forecasting results are shown in Figure 5-12, and the overall accuracy (E_{rms}) is 0.29% for the charging power, and is 0.09% for the discharging power.



a. Charging rate estimation



b. Discharging rate estimation



c. State of charge estimation

Figure 5-12. Battery model validation results

Battery state of charge is an internal variable which is not easily measurement in the real field. Fortunately, when this battery model connects with the building models, only the charging and discharging rate will affect the building operation, and the SOC as an internal variable has no effect to the overall building cluster models.

5.6 Conclusion and Future Work

In this chapter, grey box models for alternative energy generation and storage devices, as well as for whole building simulation have been developed and validated using simplified physics models based on the building emulator simulation results. Over 90% accuracy has been achieved in all these models against the detailed physics model, and the simulation speed has been improved dramatically.

6. CHAPTER 6 On-line Building Cluster Model Integration and Calibration

6.1 Background and Motivation

As introduced in section 1.2.5, building energy forecasting models' accuracy may decrease after a while when the system operation condition changes. Therefore, applying the data fusion techniques to calibrate the model can improve the model performance. The common data fusion methods are based on state estimation method, such as Kalman filter and its nonlinear formations, such as Extended Kalman Filter and Particle Filter. Therefore the background information of state estimation methods, and potential data fusion techniques for building energy forecasting are introduced in this chapter firstly. Following this introduction, a feasibility study of applying Kalman filter on the system identification model is demonstrated in section 6.3.

6.1.1 State Estimation

The general linear discrete-time system state space model can be expressed as stochastic difference equations:

$$x_k = A_{k-1}x_{k-1} + B_{k-1}u_{k-1} + w_{k-1} \quad \text{Eq. 6.1}$$

$$y_k = C_k x_k + v_k \quad \text{Eq. 6.2}$$

Where x is the state, u is the input, and y is the measurement of the system. Process noise w and measurement v are white, zero-mean, uncorrelated random noises. Described as a generic state-space model, the stochastic state estimation problem can be illustrated by a graphical model in Figure 6-1.

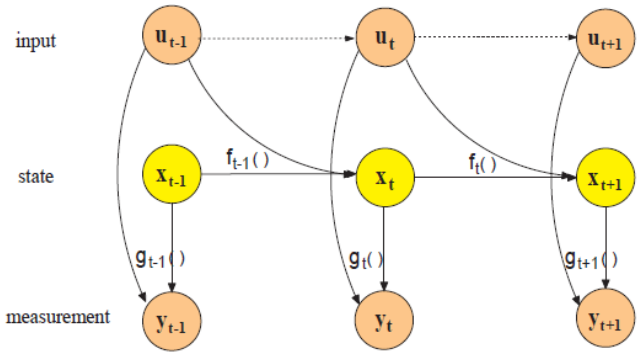


Figure 6-1. A graphical model of generic state-space model [237]

6.1.2 State Estimation Methods

6.1.2.1 Kalman Filter

The Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) means to estimate the state of a process, in a way that minimizes the mean of the squared error. The filter is very powerful in several aspects: it supports estimations of past, present, and even future states, and it can do so even when the precise nature of the modeled system is unknown. The concept of Kalman filter is to estimate the state of a dynamic system by using a feedback process: a state space filter estimates the process and receives feedback from noise measurements (Figure 6-2).

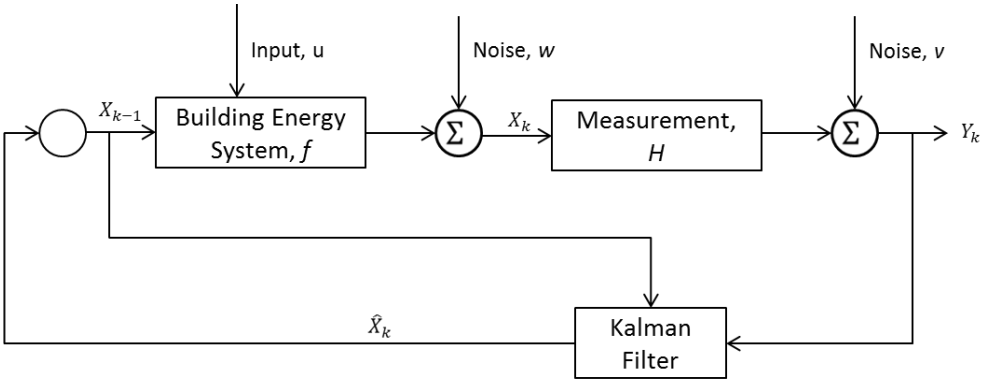


Figure 6-2. Kalman filter state estimation

The state space model will forecast forward (discrete time) from the current state and process covariance to get a priori estimation of the state at next time step. The noisy measurements will be incorporated in the system model to update the posteriori forecasting (Figure 6-3).

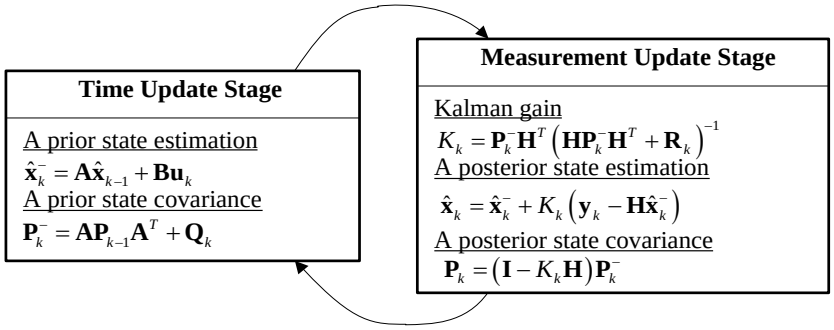


Figure 6-3. Kalman filter operation [238]

6.1.2.2 Extended Kalman Filter

As described above, Kalman filter address the state estimation problem in a linear stochastic system. Extended Kalman filter is a nonlinear version of Kalman filter. The general idea is to linearize the nonlinear system model first using the current state and covariance, and then the general Kalman filter is applied on the linear estimation model. Extended Kalman filter has also been used in some building energy forecasting studies.

An integrated 3R2C and EKF (Extended Kalman Filter) model was developed to estimate the building energy consumption in [162]. In this work, an EKF was used to estimate the state vector *X* using real sensor measurement data. The estimated load matched the EnergyPlus results within 10% at 93% of the time. This RC-EKF approach has also been tried in [163]. In this study, a self-adaptive thermal building model was developed based on a 1R1C model and an EKF. Existing Kalman filter studies show good potential for this technique, when combined with other techniques, to further improve the building energy model accuracy and robustness.

6.2 Framework of System Identification and Data Fusion

Data fusion is a process of integrating multiple data from different sources representing the same real world system into a more accurate and robust representation. In this section, a framework (Figure 6-4) of applying data fusion techniques on system identification model is developedd.

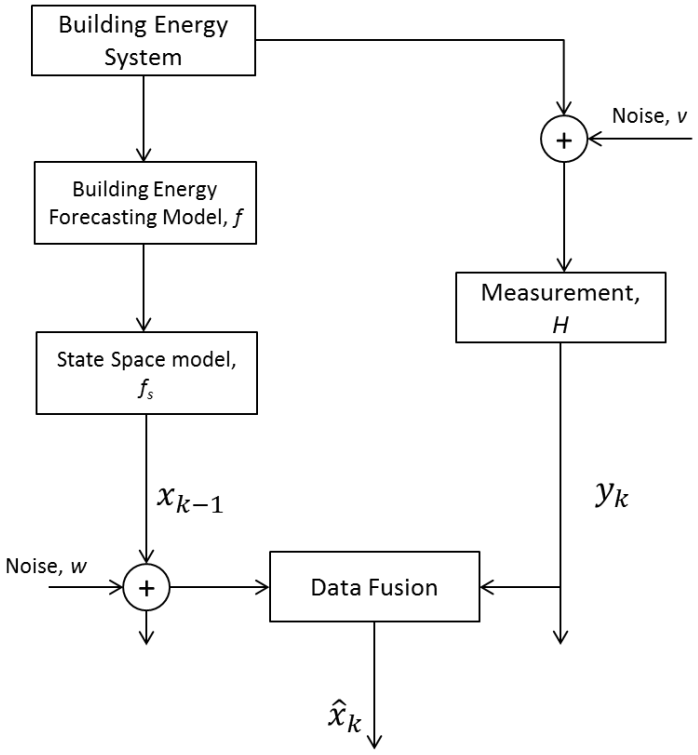


Figure 6-4. General data fusion procedure for model development

6.2.1 Building Energy State Space Model Development

The system identification model developed previously is stored as a set of Markov parameters. In order to utilize Kalman filter into this model, however, state space model format is needed. The general linear discrete-time system state space model can be expressed as stochastic difference equations:

$$x_k = A_{k-1}x_{k-1} + B_{k-1}u_{k-1} + w_{k-1} \quad \text{Eq. 6.3}$$

$$y_k = C_k x_k + v_k \quad \text{Eq. 6.4}$$

Where x is the state, u is the input, and y is the measurement of the system. Process noise w and measurement v are white, zero-mean, uncorrelated random noises.

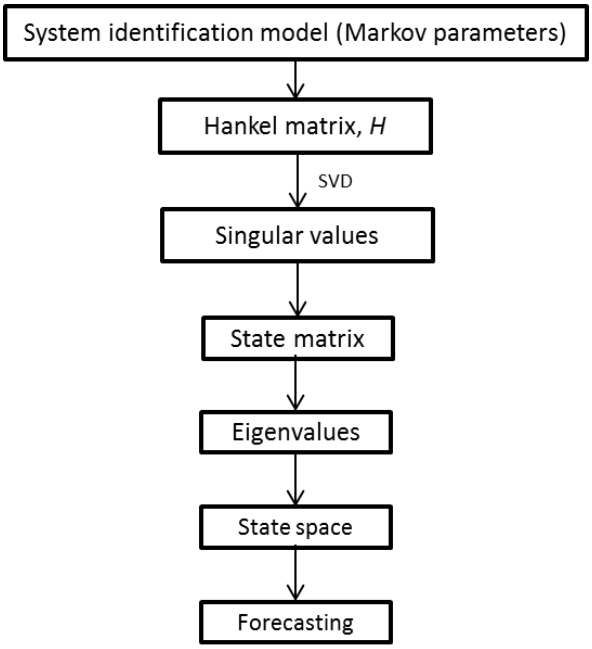


Figure 6-5. State space model through ERA

In order to transit Markov parameters into a state space format dynamic model, an eigensystem realization algorithm (ERA) developed by Juang and Pappa [239] is applied. The overall procedure of state space model formation through ERA is illustrated in Figure 6-5. The details of Hankel matrix generation, state matrix formation and system order determining, etc. are introduced as follows.

From Eq. 6.3 and Eq. 6.4, the Markov parameters can be expressed as:

$$MP_k = CA^{k-1}B$$

All these Markov parameters MP_1, MP_2, MP_3, … are calculated in the frequency response model (Figure 3-8) without explicit knowledge of the system matrices $A, B, C,$ and D. ERA is started from a system realization matrix: Hankel matrix, which is composed of the Markov parameters:

$$H(k-1) = \begin{bmatrix} MP_k & MP_{k+1} & \cdots & MP_{k+\beta-1} \\ MP_{k+1} & MP_{k+2} & \cdots & MP_{k+\beta} \\ \vdots & \vdots & \ddots & \vdots \\ MP_{k+\alpha-1} & MP_{k+\alpha} & \cdots & MP_{k+\alpha+\beta-2} \end{bmatrix}_{\alpha m \times \beta r}$$

Where α and β are determined by the number of Markov parameters from the system identification model, which should be greater than system order n.

Since building energy model is a dynamic model, whose system order could be changing according to the system operation situation (for example during starting up and shutting down period). Therefore, singular value decomposition (SVD) was then applied on $H(k)$ to determine the system order and state space model parameters:

$$H(k) = R\Sigma S^T$$

Where Σ contains the singular values of H on its diagonal, R and S are orthonormal matrices containing their corresponding singular vectors. The rank of Σ is the order of this system model.

$$\Sigma = \begin{bmatrix} \Sigma_n & 0 \\ 0 & 0 \end{bmatrix} \text{ with } \Sigma_n = \text{diag } [\sigma_1, \sigma_2 \ldots \sigma_n]$$

Therefore, the parameter s of the state space mode can be obtained from the following equations:

$$\hat{A}_k = \Sigma_n^{-\frac{1}{2}} R_n^T H(k) S_n \Sigma_n^{-\frac{1}{2}}$$

$$\hat{B}_k = \Sigma_n^{\frac{1}{2}} S_n^T E_r$$

$$\hat{C}_k = E_m^T R_n \Sigma_n^{\frac{1}{2}}$$

Where $E_r^T = [I_r\ O_r \dots O_r]$ and $E_m^T = [I_m\ O_m \dots O_m]$, where I is identical matrix, r is the number of inputs and m is the number of outputs. R_n and S_n are the first n columns if matrices R and S, respectively. $\hat{A}_k$, $\hat{B}_k$ and $\hat{C}_k$ are the estimation of state space model parameters: A_k, B_k and C_k.

6.2.2 Kalman Filter Implementation on State Space Model

Recall the state space equations in Eq. 6.3 and Eq. 6.4. Noises w_k and v_k have known covariance matrices Q_k and R_k, respectively. Q_k is determined by the discrepancy of the forecasted and real energy consumption. R_k is determined by the measurement noise. In this study, simulation results from the validated EnergyPlus model were used as the ground truth (real energy consumption), and measurement error stand deviation was chosen as 10%:

$$w_k \sim N(0, Q_k)$$

$$v_k \sim N(0, R_k)$$

The concept of Kalman filter is to estimate the state of a dynamic system by using a feedback process: a state space filter estimates the process and receives feedback from noise measurements (Figure 6-2). The state space model will forecast forward (discrete time) from the current state and process covariance Q_k to get a priori estimation of the state at next time step. The noisy measurements will be incorporated in the system model to update the posteriori forecasting (Figure 6-3). In this study, in order to represent the real situation and apply data fusion techniques, a white measurement noise is added to the output from EnergyPlus test bed:

$$y_k = EP_k + EP_k * \sqrt{\sigma_k} * N_{rand}$$

Where,$y_k$is energy consumption measurement, EP_k is the noise-free energy consumption from EnergyPlus, $EP_k * \sqrt{\sigma_k} * N_{rand}$ is measurement noise, σ is the measurement error variance, and N_{rand} is a normal random number between -1 and 1.

In this study, a system order auto checking model has been developed before state space model generation and Kalman filter implementation. In the state space model, state vector $X_k = [E_k\ E_{k-1} \dots E_{k-n}]$ (n is system order), measurement vector $Y_k = [E_k]$, and input vector contains all the input variables in the system identification model, as described in Table 3-1.

Follow the Kalman filter implementation procedure discussed before, the process and measurement covariance are determined as:

$$Q_k = I(n) * \sigma_k$$

$$R_k = (EP_k * \sqrt{\sigma_k} * N_{rand})^2$$

where $I(n)$ is a $n \times n$ identity matrix, σ_k is calculated from the state space model forecasting error. Then Kalman filter is applied to update the state space model and to improve the forecasting accuracy every time step.

6.3 Feasibility Study

The same small size commercial building introduced in section 3.5.1 is used in this chapter for building energy on-line forecasting. As introduced early, the validated EnergyPlus simulation test bed is used as a “real” building to provide training data and validation data for both heating and cooling seasons. The overall procedure of the generation and validation of the on-line energy estimation model is shown in Figure 6-6. First, frequency response function is applied to create a system identification model to forecast the energy consumption, and then the ERA method is used to reformat the system identification model to be a state space model, which is also used to

forecast the energy consumption. At the last step, Kalman filter is implemented to update the energy consumption forecasting based on the "real" measurements. There are two sets of validation data, one is used to validate the system identification model and state space model and the other is used as "real" measurements to update and validate the Kalman filter on-line model. The difference between these two sets of validation data is that "real" measurements are corrupted with measurement noises. The methods used in each step have been introduced in previous sections, as the results will be discussed in following ones.

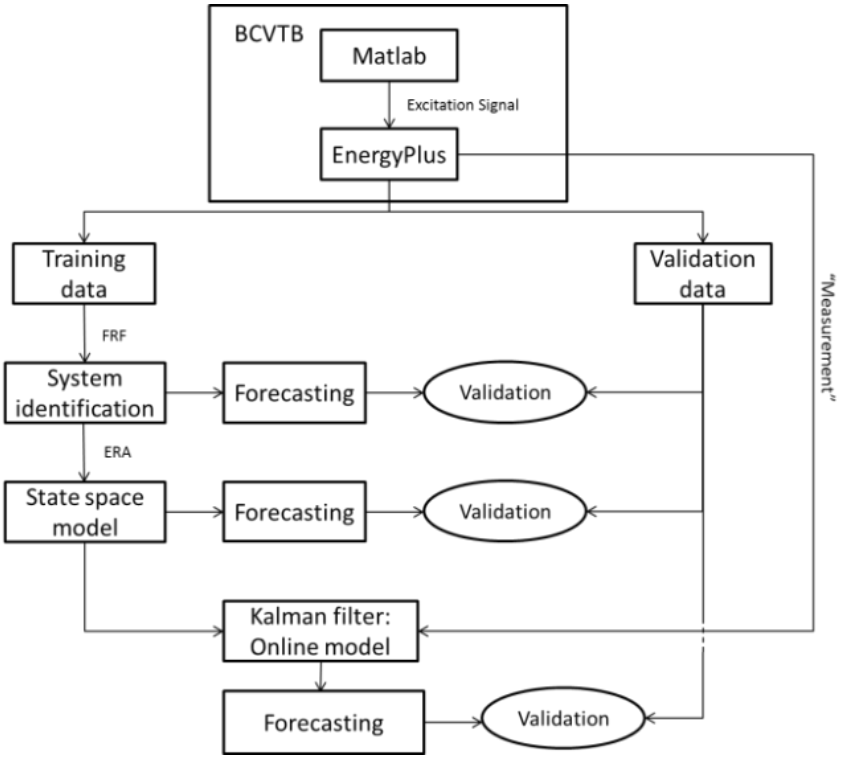
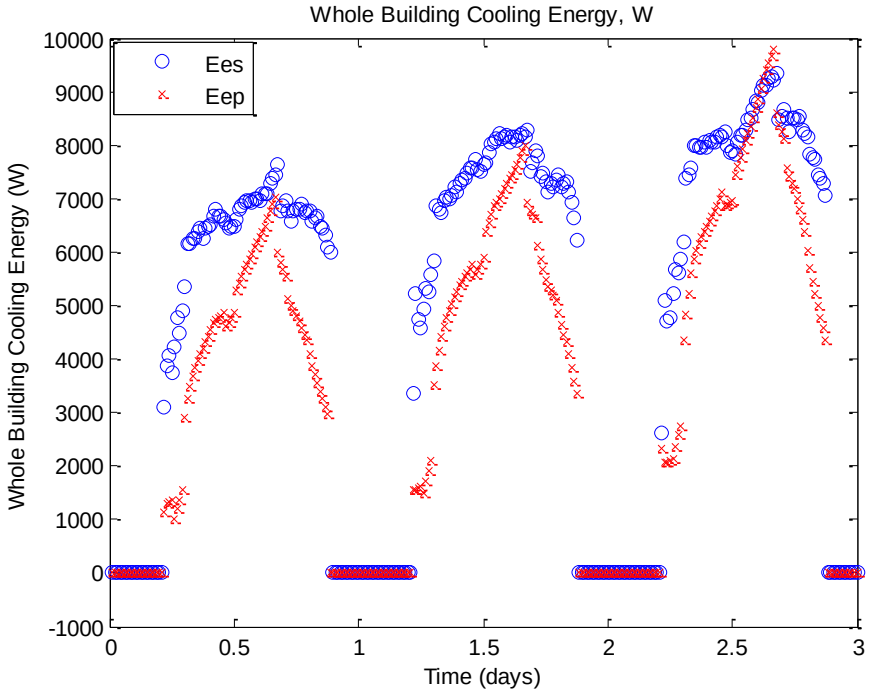


Figure 6-6. On-line model development procedure

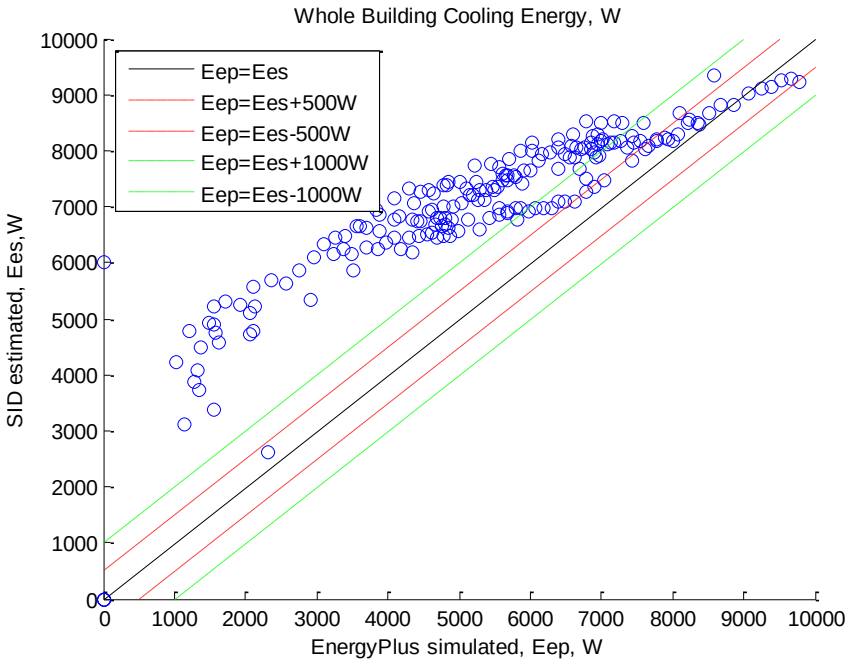
By using the ERA method that discussed before, the system identification model is then regenerated to be a state space model based on the reformation of Markov parameters. The order of the state space model varies from 4 to 36 in this study, due to the order of dynamics of the building energy system at different operation situations.

6.3.1 Energy Forecasting Speed

The state space model sacrificed some accuracy due to the truncation of the Markov parameter according to the system order. As it is shown in Figure 6-7, the error between forecasting results from state space model (*Ees*) and EnergyPlus simulation results (*Eep*) are much larger than that in Figure 3-10. The overall R^2 is just 0.71. However, the state space model is still able to follow the trend of the EnergyPlus estimation results, and it does not under estimate the cooling energy at noon of the third day when cooling energy consumption is very high. In the next step, Kalman filter based data fusion techniques will then be applied to improve the forecasting accuracy and robustness. In Figure 6-8, the energy forecasting results are plotted. Comparing to the results in Figure 6-7, the forecasting accuracy has been significantly improved.

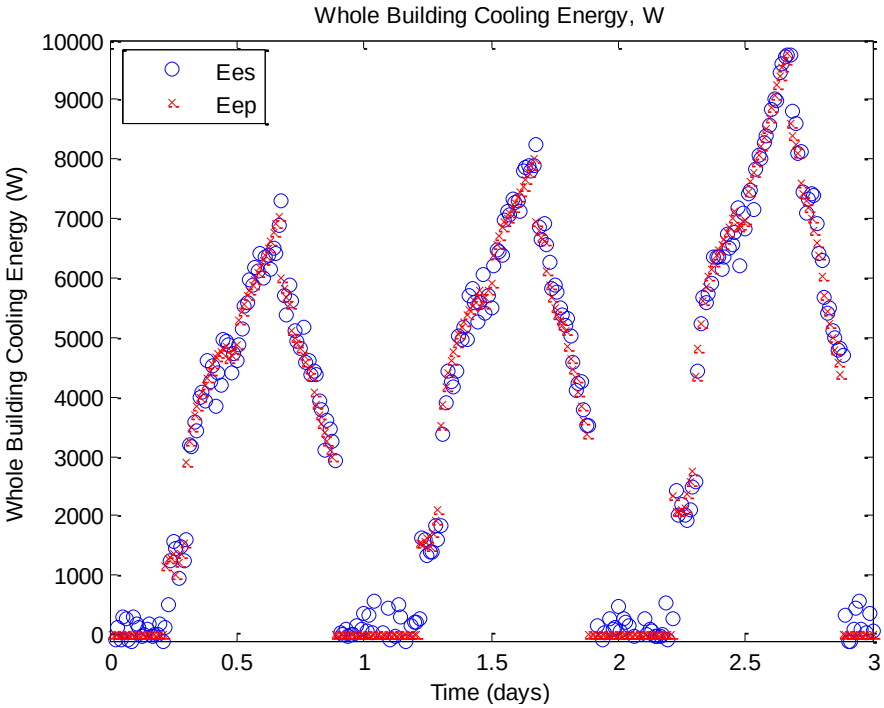


a) Building energy forecasting results

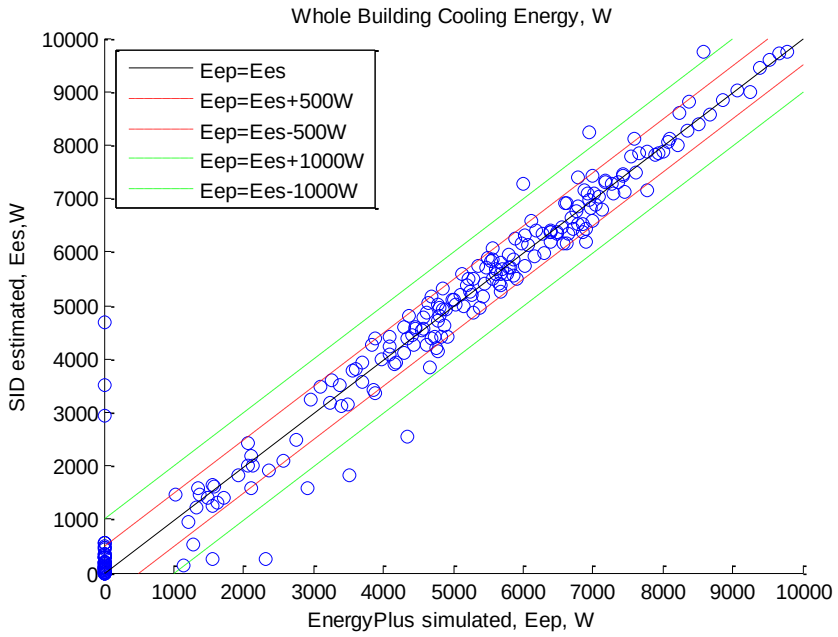


b) Building forecasting error analysis

Figure 6-7. State space model forecasting results.



a) Building energy forecasting results



b) Building forecasting error analysis

Figure 6-8. Kalman Filter on-line forecasting results

Above all the forecasting accuracy of these three models is summarized in Table 6-1, using the two evaluation indexes, R^2 and N_{rms} introduced in section 3.5.2. Kalman filter model with the on-line updating based on the real measurement has the highest accuracy, while the state space model has the lowest accuracy, because it truncated the Markov parameters based on the system order.

Table 6-1. Model forecasting accuracy comparison

Model	System Identification	State Space	Kalman Filter
R^2	0.95	0.71	0.97
$\boldsymbol{N_{rms}}$	14.7%	21.7%	5.8%
Speed (s)	0.0089	0.03	0.02

6.3.2 Energy Forecasting Speed

Simulation speed is another crucial factor for MPC in practical use. The forecasting speeds of these three models are also summarized in Table 6-1. The system identification model training time is around 48 second, as show in Table 3-4. Fortunately, in any MPC model, energy forecasting model doesn't need to be trained every time step. In this study, this model is just trained once for three days' forecasting, once the Markov parameters were calculated from the training data, they were saved and used by the on-line MPC model every time step for operation optimization.

6.4 Building Cluster Model Calibration Refinement

After the success of the feasibility study of the building energy forecasting model calibration, the state space model reconstruction and data fusion approaches are refined and improved in two directions:

1. System order dynamic selection. In the updated state space model reconstruction approach, the system order is selected at each time step using ERA method. This dynamic system order selection method improved the model performance because building energy system performance differently and has different system characteristics at different operation region, such as starting up and shutting down period.
2. Model performance under noisy condition. Similar to the small building calibration study in section 6.3, the process uncertainties for inputs variables: T_{out}, Q_{dir}, Q_{dif}, R_{in}, and V_{oa} are considered as:

$$X_i = \tilde{X}_i + \varepsilon_i \qquad \text{Eq. 6.5}$$

Where, i stands for T_{out}, Q_{dir}, Q_{dif}, R_{in} and V_{oa}; X is the noisy value, $\tilde{X}$ is the true value, and ε is the noise of each variable. In this study, ε is assumed to follow a standard normal distribution within 5%.

6.4.1 Small Building SID Model Calibration Results

Based on the previous data fusion feasibility study and the refining studies, the forecasting results for the small building are improved, as shown in Figure 6-9. The detailed forecasting result statistics are summarized in Table 6-2. The adapted SID model is able to achieve around 95% accuracy under the noisy free conditions for small building cooling energy forecasting, while in the noisy conditions the stat e space model is only able to get around 87% accuracy, and at last the Kalman filter is able to improve the forecasting accuracy to above 95%.

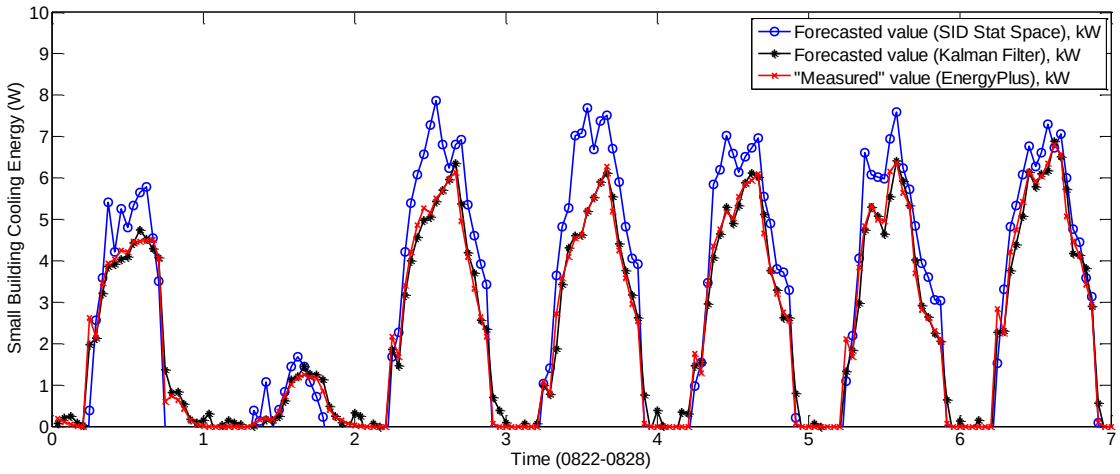


Figure 6-9. Small building energy forecasting model calibration results

Table 6-2. Model forecasting accuracy comparison: small building

Model	R^2	RMSE, kW	NRMSE
SID (noise free)	0.95	0.98	14.7%
State Space (noisy)	0.87	2.67	22.1%
Kalman Filter	0.98	0.36	3.12%

6.4.2 Medium Building SID Model Calibration Results

Using the same SID model, state space model, and Kalman filter model, the energy consumption for the medium building from 0811 to 0817 and 0822 to 0828 are forecasted under noisy condition. Figure 6-10 compared forecasting results from state space model and Kalman filter against those from EnergyPlus. Similar to the situation in the small building, the sate space model result has larger fluctuation than that of the Kalman filter model, and Kalman filter model also improved the overall forecasting accuracy from 83% to 97% by incorporating energy consumption measurements (Table 6-3). In this study, the weather conditions in forecasting period are very close to those in the training period. The outdoor temperature range in the training period is 20.1 ºC – 35.6 ºC, and that in the forecasting period is 18.6 ºC - 35.1 ºC. Even though the temperature in the forecasting period at some time is lower than the minimum temperature of the training period, the number of this situation is very small and the all happened at night between 11 pm to 5 am when the building temperature setpoint is very loose. Therefore the accuracy of the adapted SID model is very good, and the improvement of the Kalman filter is not very obvious. On the other hand, the state space model overestimated the energy forecasting at the last day. This is because the last day is Sunday, when building temperature setpoints are 26.7 ºC (80.1 ºF) for cooling setpoint and 15.6 ºC (60.1 ºF) for heating setpoint, and the energy consumption is much lower. In this study, the energy forecasting model is developed upon 1 week building operation data, so there is only one Sunday comparing to five week days, and the energy forecasting models is developed upon more data under week day operation. Therefore the SID state space model trends to overestimate the energy consumption for the weekends. In order to solve this problem, more weekend operation data should be included for model training, or different models for weekdays and weekends are necessary.

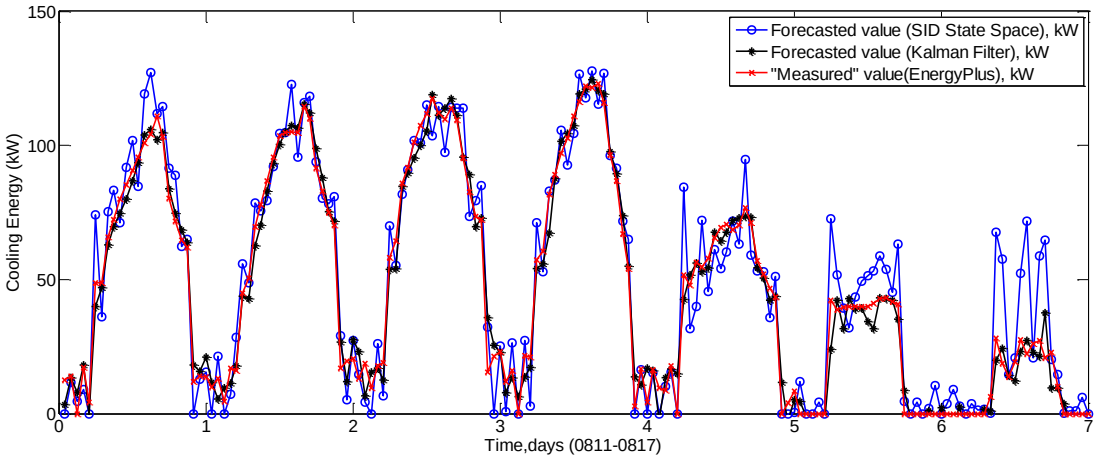


Figure 6-10. Medium building energy forecasting model calibration results

Table 6-3. Model forecasting accuracy comparison: study I

Model	**R^2**	**RMSE, kW**	**NRMSE**
SID (noise free)	0.96	4.90	4%
State Space (noisy)	0.83	16.81	14%
Kalman Filter	0.97	3.46	3%

In order to show the effectiveness of the Kalman filter, another study is conducted here, using the SID model training upon the building operation data from 08/01-08/07 to forecast the energy consumption from 08/22-08/28. The energy forecasting results in this study are plotted in Figure 6-11. Similar to the results in study I, the SID state space model has large fluctuations in the forecasting results and overestimated the energy consumption, especially in the weekends and the Kalman Filter results are very close to the "measured" value. The detailed performances of these three models are summarized in Table 6-4. The SID state space model has the lowest accuracy with R^2 as 0.76, and the Kalman filter also achieved over 90% accuracy under noisy conditions.

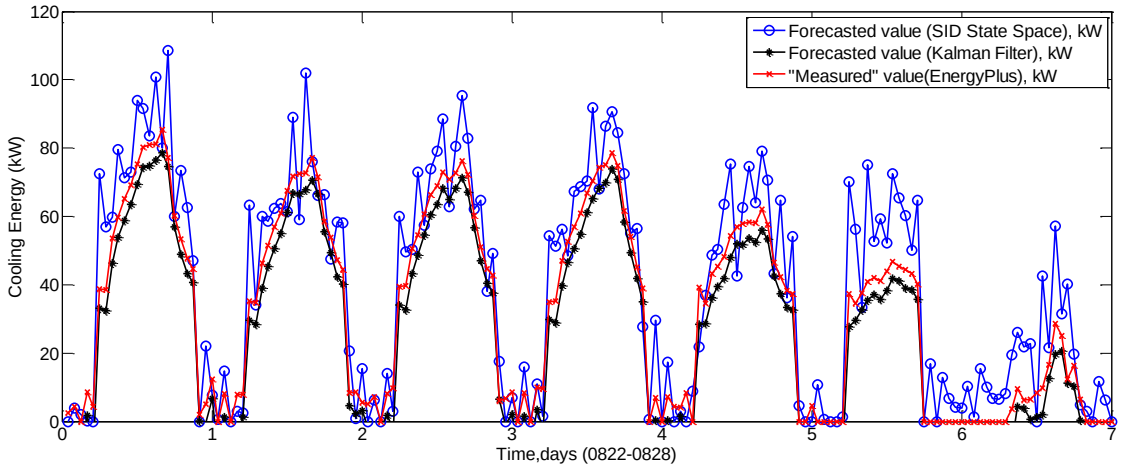


Figure 6-11. Medium building energy forecasting model calibration results: study II

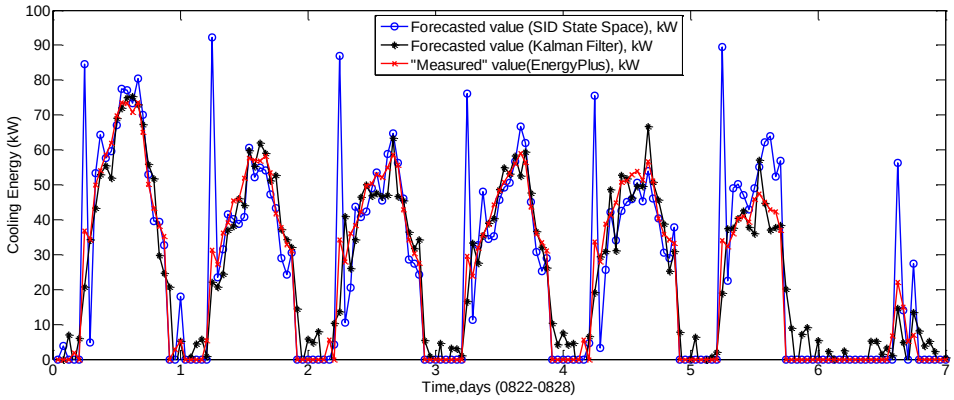
Table 6-4. Model forecasting accuracy comparison: study II

Model	**R^2**	**RMSE, kW**	**NRMSE**
SID (noise free)	0.92	5.92	8%
State Space (noisy)	0.76	16.81	18%
Kalman Filter	0.92	6.64	9%

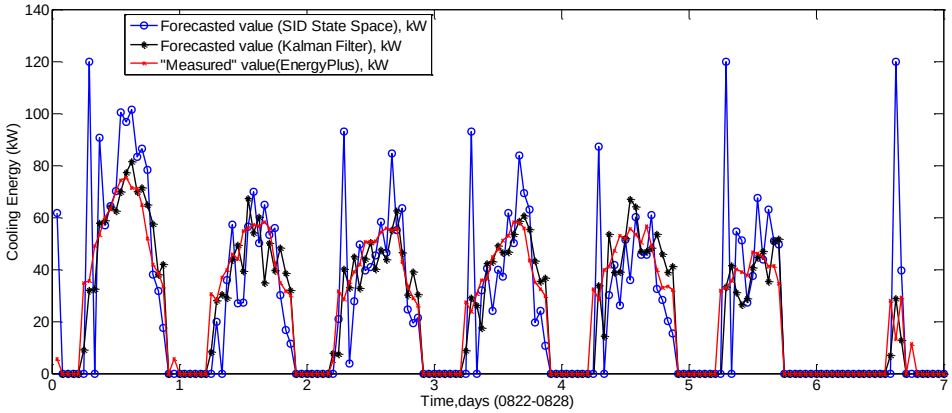
6.4.3 Forecasting and Data Fusion Time Interval

In all the previous studies, the forecasting and data fusion time step are 15 minutes. In this section, 30 and 60 minutes forecasting time intervals will be studied. Except the time interval, other conditions, such the targeting building, forecasting period, and the weather conditions, etc. are all remained the same. The same energy forecasting model developed for 15 minutes time step forecasting in section 6.4.1 and 6.4.2 will be used here for different time step energy forecasting during 0822-0828. The following figure (Figure 6-12) is the energy forecasting results for medium building energy forecasting. The upper plot Figure 6-12a) is the forecasting results with 30 minutes forecasting interval, and the lower one (Figure 6-12b) is forecasting results with 60 minutes forecasting interval. In these two plots, both SID state space model and Kalman Filter model are under 5% noisy condition. These two plots clearly show that the SID state space model

(blue line) is still able to capture the trend of the energy consumption and can achieve good accuracy in the forecasting, but the fluctuation of the results is larger than that of Kalman Filter (black).



a) 30 minutes time interval



b) 60 minutes time interval

Figure 6-12. Medium building energy forecasting model calibration results:

The detailed forecasting accuracy statistics for different models with all these 3 intervals are summarized in Table 6-5, which clearly shows that the models with shorter time interval have higher accuracy, Kalman filter models have the highest accuracy comparing with SID model and State space model.

Table 6-5. Model forecasting accuracy comparison

Forecasting Interval, min	Model	R^2	RMSE, kW	NRMSE
15	SID (noise free)	0.92	5.92	8%
	State Space (noisy)	0.76	16.81	18%
	Kalman Filter	0.92	6.64	9%
30	SID (noise free)	0.82	10.9	14.0%
	State Space (noisy)	0.72	12.5	15.3%
	Kalman Filter	0.91	6.9	9%
60	SID (noise free)	0.76	12.9	15.0%
	State Space (noisy)	0.69	13.4	18.0%
	Kalman Filter	0.86	8.7	11.2%

6.5 Building Cluster Model Integration

One of the major objectives of this thesis is to develop a high fidelity building cluster model that can be used for on-line model based control and operation optimization. From the studies described above, forecasting models for different size of building, one PV panel energy generation system, one battery and one ice tank energy storage systems have been developed and validated individually. In this section, they are integrated together forming an on-line building cluster simulation model. The detailed declarations of each model in this on-line model are summarized in Table 6-6. Upon all these models, there will be an overall operation and decision model which will generate all the control variables for each component model.

Table 6-6. Building cluster model declaration

	Input (output of decision model)	Output
Building	Temperature setpoint, Power from battery, power from PV, power from grid	Energy consumption (cooling and non-cooling)
Battery	Battery state	Charging power: power from grid, power from PV Discharging power: power to building, State of charge
PV	PV state	Power generation: Power to building Power to grid Power to battery
Ice tank	Ice tank state	Charging: from grid Discharging power State of charge
Decision	Outputs from all the component models, tariff (TOU)	all inputs of each component model

6.6 Conclusion and Future Work

An approach for building energy forecasting model on-line calibration using data fusion techniques is introduced in this chapter. Based on the system identification preliminary study in Chapter 3, eigensystem realization algorithm is used to reformat the system identification model to be state space model from Markov parameters. Finally, Kalman filter was applied to the state space model to update the energy forecasting and improve the forecasting accuracy and robustness. 15, 30 and 60 minutes updating time interval have been tested. This on-line building energy model achieved over 95% forecasting accuracy for cooling energy consumption in a small building, and achieved over 90% accuracy in a medium building studied here.

The future work on the data fusion part will focus on the real field application and validation. The details about the real field experiment will be discussed in Chapter 7.

7. CHAPTER 7 Building Energy Forecasting Methodology Verification

7.1 Background and Objectives

In this chapter, the novel building energy forecasting method developed in pervious chapters is applied in a real building to verify the performance. The preliminary studies show great promises in terms of their accuracy, robustness, and cost-effectiveness using simulated data from the cluster emulator testbed. In order to examine and validate the performance of developed models under real world conditions, experiments in a real building are desired.

The overall objectives of this experimental study are to design and conduct experiments to evaluate the system identification methodology described in Sec XX for building on-line energy forecasting, and document all testing procedures and results. This includes:

1. Systematically evaluate the nonlinearity and response time of the test building, which includes the test building envelope and HVAC equipment;
2. Design and apply system excitation strategies based on the nonlinearity test results to generate operation data for system identification model development and validation;
3. Develop and validate the system identification model for building cooling energy forecasting.

7.2 Overall Experiment Scope and Plan

The experiment scope and detailed plan used to fulfil the objectives of this experimental study are described in this section.

7.2.1 General Description of Experiment Facility

All of the experiments have been conducted at the Energy Resources Station (ERS) of Iowa Energy Center. The ERS is a small size commercial building with two full-scale commercial

HVAC systems side by side with identical thermal loadings and weather conditions, which allows testing and comparing two different operation schemes. A schematic diagram of the floor plan for the ERS is shown in Figure 7-1. The facility is equipped with three VAV air handling units (AHU). AHU-1 serves the common areas of the building. The remaining two AHUs serve the A- and B-Test Systems. AHU- A and B are identical, with each AHU serving four zones. Therefore, in this experimental study, different building HVAC systems and operation schemes can be used in parallel in these two systems. In each room of this testing facility, there are two different lighting stages, and two base board heater stages.

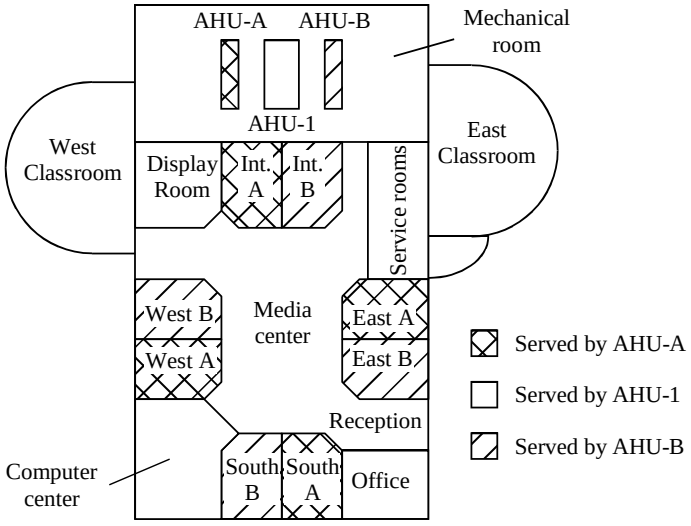


Figure 7-1. Floor plan of energy resource station at Iowa Energy Center

7.2.2 General Experiment Scope

Chapters 3 to 5 introduced the building energy forecasting model development, using system identification and simplified physics methods. Based on this preliminary study and considering the objectives of the experiment, this experiment has been divided in three sub-tasks: 1) building energy system nonlinearity test for system identification prior information collection; 2) building

energy system excitation for system identification model development and validation; 3) building online energy forecasting model development and validation.

7.3 System Nonlinearity and Excitation Methodology: Experiment Design

Similar to the system excitation in EnergyPlus model, the excitation signals will also be applied in real building energy systems. Therefore in this building energy system identification experiment test, building temperature setpoints and internal equipment operation schedules generated in simulation models in section 4.4.3 have been applied into the building automation system to control the building. The building system's nonlinearity is calculated by evaluating the system inputs and outputs. Several important parameters determined in simulation study are adapted and adjusted according to the real building performance, such as excitation signal variations, signal injection intervals.

Building heating and cooling load and energy consumption are measured or calculated at each time step under different operation strategies. Similar to the system identification model development in Chapter 3, a new building energy forecasting model is then be developed and verified based on the real field system excitation. The experiment is conducted from 08/24/2015-09/04/2015, and 09/15/2015-09/20/2015. In the first round experiment, the system nonlinearity test is conducted at first three days, followed by the one day system response time test. Then 7 days system excitation and 3 days normal operation test are conducted. The detailed testing operation plan can be found in Appendix C.

7.3.1 Real Field System Nonlinearity Test

Similar to the system nonlinearity study using simulated data (Chapter 4), pre-determined system nonlinearity test signals, such as temperature setpoints (heating and cooling) and equipment schedules, are applied to the real building systems. The testing signals are pre-tested

using simulation models and are updated every 30 minutes. In this real field study, two baseboard heaters have been used to represent the internal equipment. Each base board heat has a capacity of 900 W. Besides the base board heaters, lighting system is also included in this system nonlinearity test study. At the ERS, there are two stages of lighting system operation. Therefore there are 7 stages for the internal equipment schedule:

1. Stage zero: all off
2. Stage one: lighting stage 1 on only, baseboard heater off
3. Stage two: lighting stage 2 on only, baseboard heater off
4. Stage three: lighting stage 1 and 2 on, baseboard heater off
5. Stage four: Baseboard heater stage 1 on, lighting system off
6. Stage Five: Base board 1 and 2 on, lighting system off
7. Stage Six: all on

The detailed temperature setpoint and equipment stages settings during the nonlinearity test period are provided in Appendix C.

7.3.2 Real Building System Response Time Test

In this study, the building zone temperature is chosen as the measurement in this response time experiment. Since both the HVAC system capacity and the building thermal mass could affect the zone temperature response time, two tests are performed to evaluate the response time respectively. The first one is to change the zone temperature setpoint after the building zone temperature has reached a steady state, and then measure the time between the beginning of the setpoint change and when the zone temperature reaches the 95% of the temperature change. This response time reflects a combined impact from building thermal mass and HVAC system capacity. The other one is to switch off the HVAC system at night, when weather disturbances

are minimal, and measure the time that the zone temperature takes to decrease to a steady state (or nearly a steady state). Here, a nearly steady state condition is defined as less than 0.5% of the state change in 15 minutes. This second test evaluates the impact of a building's thermal mass on its state response time.

During the experiment, the temperature response time is firstly determined through historical data by checking how long it will take the ERS system to reach a new steady state after a temperature setpoint change. The temperature setpoints during test 2 need to be determined based on the outside temperature during the testing days. 5, 8, and 10 degree differences between the inside and outside temperature are maintained firstly before the system is turned off.

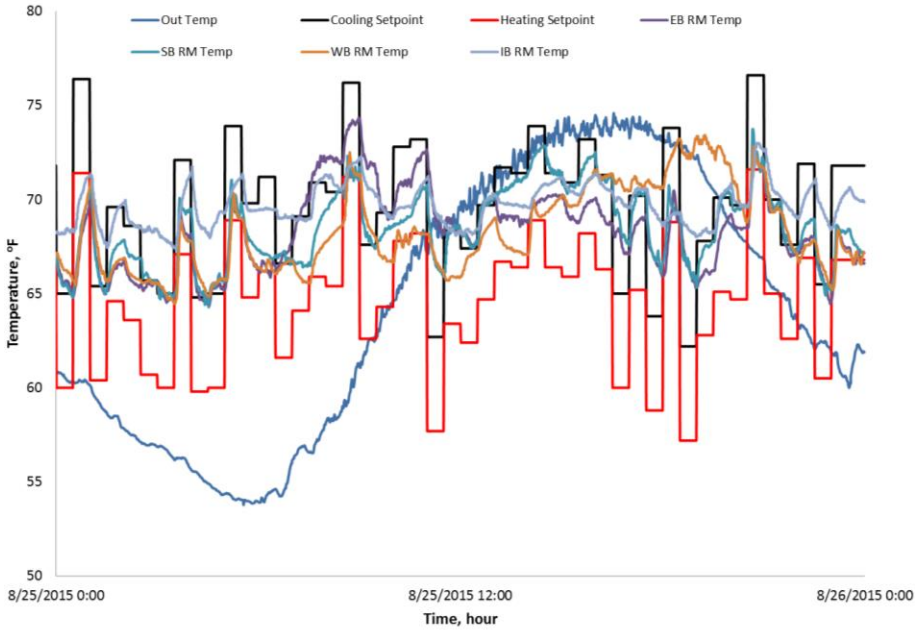
7.4 Building online energy forecasting method verification

System characteristic evaluation methodology developed in 4is applied to analyze the experiment data from the two sub-experiments (nonlinearity test and response time test). Then the test building is excited based on the excitation plan to provide training data for the SID model.

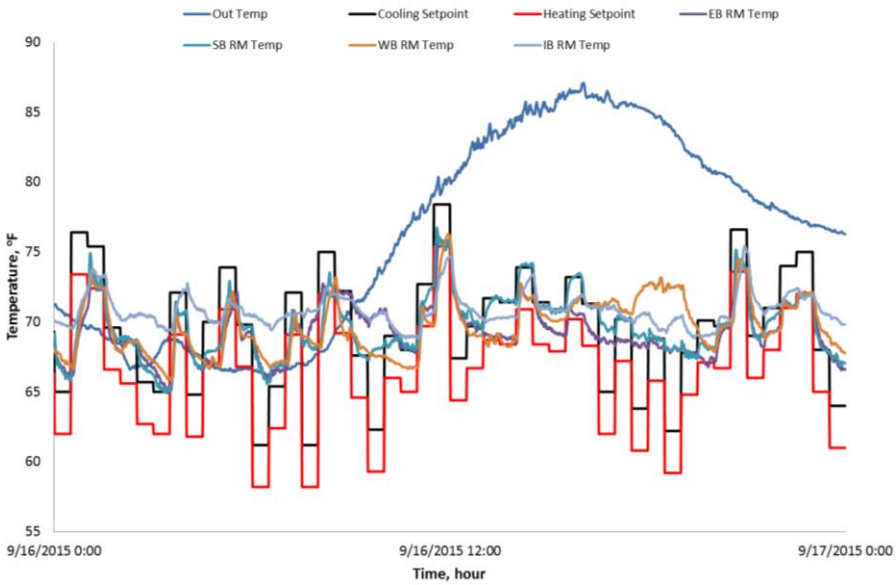
After the experiment for energy forecasting model development (training) is finished. The system is operated using normal control strategies. System operation data under normal strategies will be used for energy forecasting model validation (testing). As introduced section 7.2.1 Sec XX, there are two identical HVAC systems in this building. Hence the operation data from either system can be used for model validation.

7.4.1 Building Energy System Characteristics Test Results

Test plan as described in Section 7.2 and Appendix C is used to test the system nonlinearity and response time. The system characteristic test for nonlinearity is conducted on 09/15/2015 (test 1) and 09/16/2015 (test 2). Due to the facility schedules, interior A room is not included in this test. The room temperatures under the nonlinearity test signal are plotted in Figure 7-2.



a) System B room temperature in 0825



b) System B room temperature in 0916

Figure 7-2. System Nonlinearity test: room temperature

Using the experiment data, the system nonlinearity between each system input and output are evaluated and plotted in Figure 7-3. Similar to the simulation study for the medium building, the

system nonlinearity indexes of all six inputs are closer to one at lower frequency region. This means the system behaves more like a linear system when the input signals are at lower frequency. By checking the system input signal distribution as illustrated in Figure 7-4, all of the input signal are distributed in the lower frequency range from 1 to 0.2 h^{-1}. Therefore, under the operation used in this study, the system behaves like a linear system, which will benefit the model accuracy.

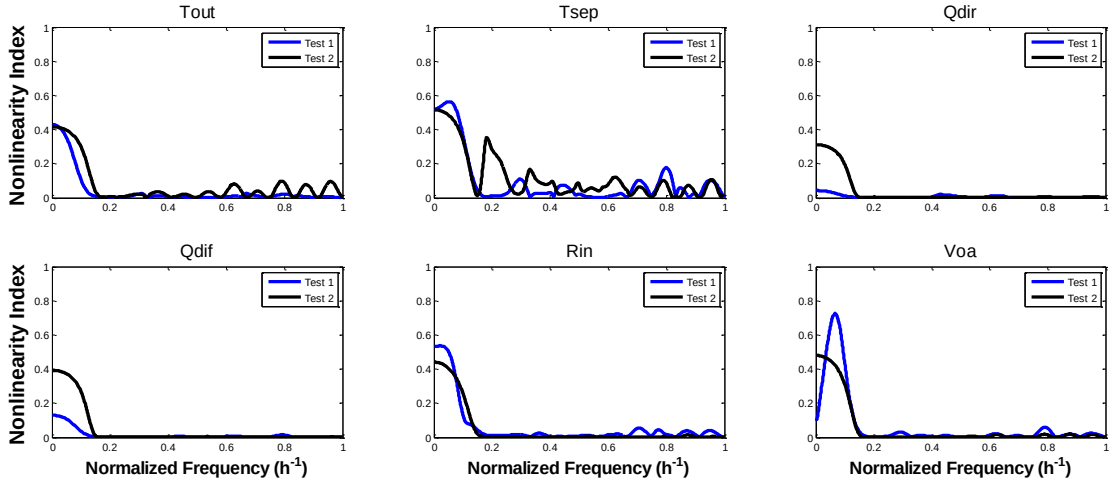


Figure 7-3. System B system nonlinearity test

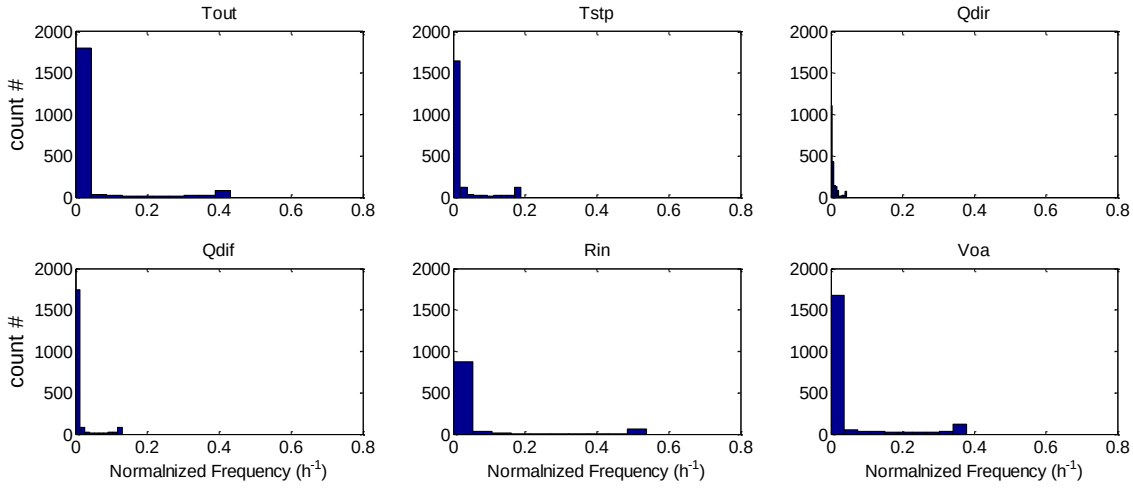


Figure 7-4. System A system input signal distribution histogram

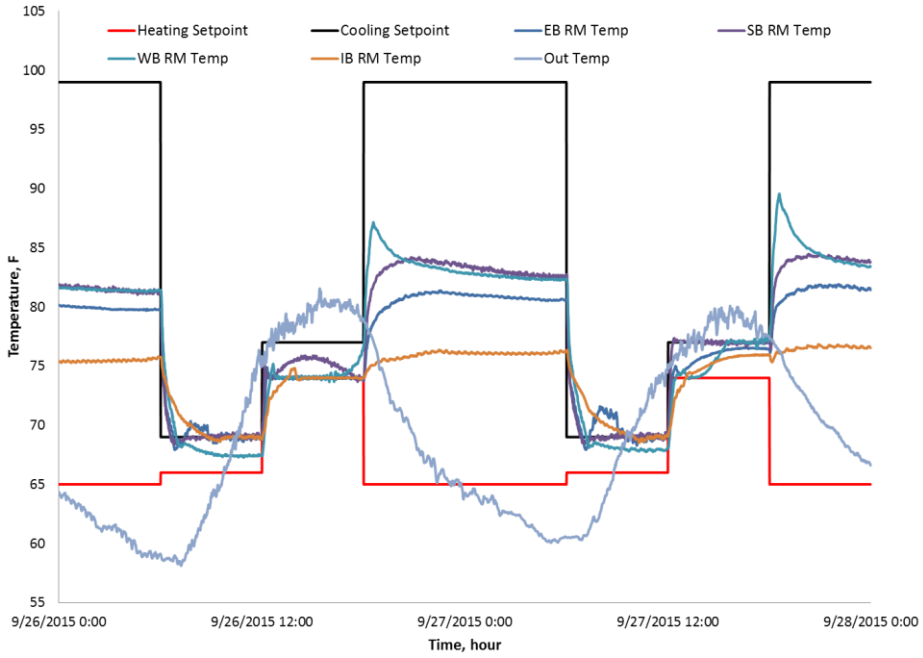
After the system nonlinearity has been evaluated, the system response time is also evaluated. The temperature setpoint settings and temperature response are plotted in Figure 7-5. The system response time are calculated following these rules:

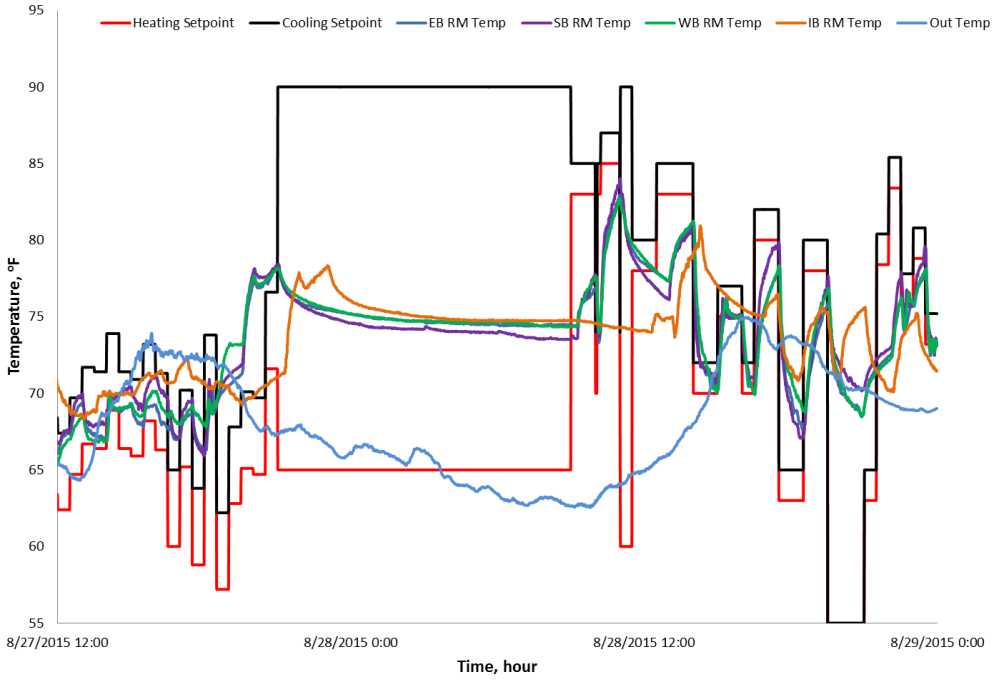
1. For system switching on to off: measuring the time for the zone temperature reaching the 95% of the final stable temperature. As the outside temperature decreases at night, the zone temperature is also decreasing when the HVAC system and internal equipment are turned off. Therefore when the temperature change is less than 0.5 degree within 30 minutes, the system is considered as stabilized;
2. For system swathing off to on: measuring the time for the zone temperature reaching to the setpoint if the system is able to control the zone temperature;
3. For temperature setpoint changing when system is on: measuring the time it takes for the zone temperature to reach a new setpoint and be stabilized.

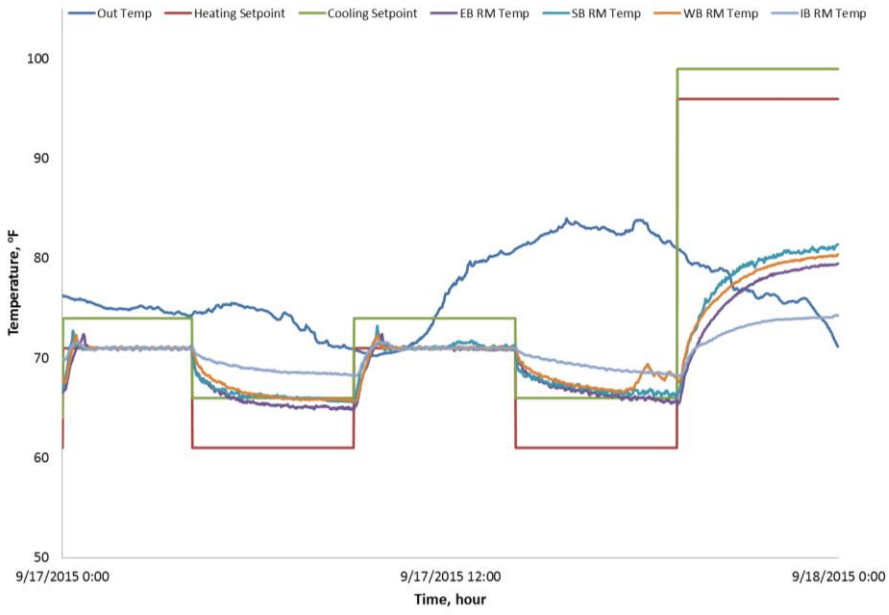
The building zone temperatures are plotted in Figure 7-5. Summarizing these plots, the system response time are evaluated and tabulated in Table 7-1. In this table, it shows that it takes around 90 minutes for the system to stablize when the system is turned off. When the system is being turned on, it takes around or less than 30 minutes for the zone temperature to stablize. For the temperature setpoints changing when the system is on, it takes about 50 minutes for the zone temperature to change from 77 to 70 ºF; and around 30 to 50 minutes from 69 to 77 ºF. Therefore, rising temperature for 10 degree typically needs 30 to 50 minutes, and decreasing temperature for 10 degree needs around 50 minutes. This response time information is then used for new system excitation signal generation.

Table 7-1. Building energy system response time test results

	Time, min			
Test (cooling setpoint, ºF)	East room	South room	West room	Interior room
77 - off	88	91	97	68
off - 77	31	21	16	-
77 - 70	51	47	-	87
70-65	46	32	-	55
69 - 77	29	32	49	117
65-70	47	67	51	56




Heating Setpoint
Cooling Setpoint
EB RM Temp
SB RM Temp
WB RM Temp
IB RM Temp
Out Temp
95
90
85
80
75
70
65
60
55
Temperature, °F
8/27/2015 12:00
8/28/2015 0:00
8/28/2015 12:00
8/29/2015 0:00
Time, hour


Out Temp
Heating Setpoint
Cooling Setpoint
EB RM Temp
SB RM Temp
WB RM Temp
IB RM Temp
100
90
80
70
60
50
Temperature, °F
9/17/2015 0:00
9/17/2015 12:00
9/18/2015 0:00
Time, hour

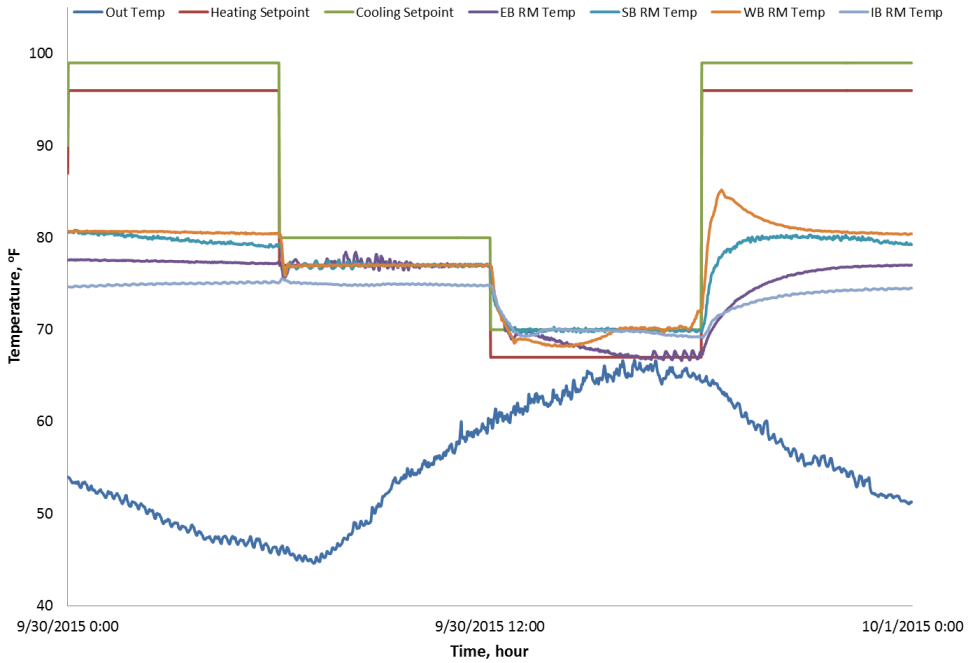


Figure 7-5. Building thermal response time test

7.4.2 Building Energy System Excitation and Identification

Based on the system nonlinearity and system response time, the system excitation signals are generated following the same procedure as introduced in section 4.6. The new signals for the temperature setpoints and equipment schedule levels are plotted in Figure 7-6. More details of the equipment schedule can be found in Appendix C.

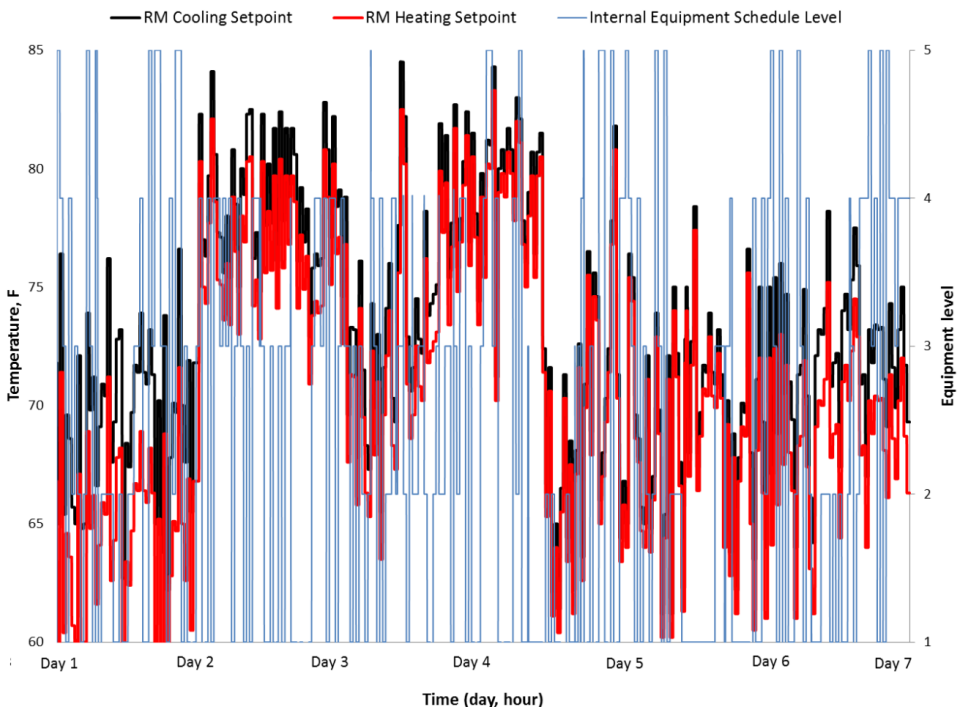


Figure 7-6. Real field system excitation signals

7.4.3 Building Energy Forecasting

Using the same frequency response function based SID approach as described in Section 3.2, the energy forecasting model is developed for the test building. The system inputs and outputs are tabulated in Table 7-2, which are the same as the SID model described in section 3.2.2.

Table 7-2. Building energy system response time test results

Variable	Variable Name	Type
$E_{chiller}$	Chiller energy consumption, kW	Output
T_{out}	Outdoor air temperature, C	Input
$T_{zone,\ stp}$	Zone temperature setpoint, C	Input
R_{in}	Equipment and occupancy heat gain, -	Input
Q_{dir}	Direct solar radiance, W/m^2	Input
Q_{dif}	Diffuse solar radiance, W/m^2	Input
Voa	Ventilation flow rate, CFM	Input

Using the same model development approach described in section 4.6, the SID model is developed using training data obtained from 0:00 am 08/28/2015 to 9:00 pm 09/01/2015 and from 9:00 pm 09/02/2015 to 12:00 am 09/05/2015. Then the developed SID model is used to forecast the building cooling energy (chiller energy) consumption in 3 different cases. The operation strategy (excitation) in the training period is illustrated in Figure 7-6. The operation strategy in the forecasting period is a typical normal building operation strategy, as show in Figure 7-7

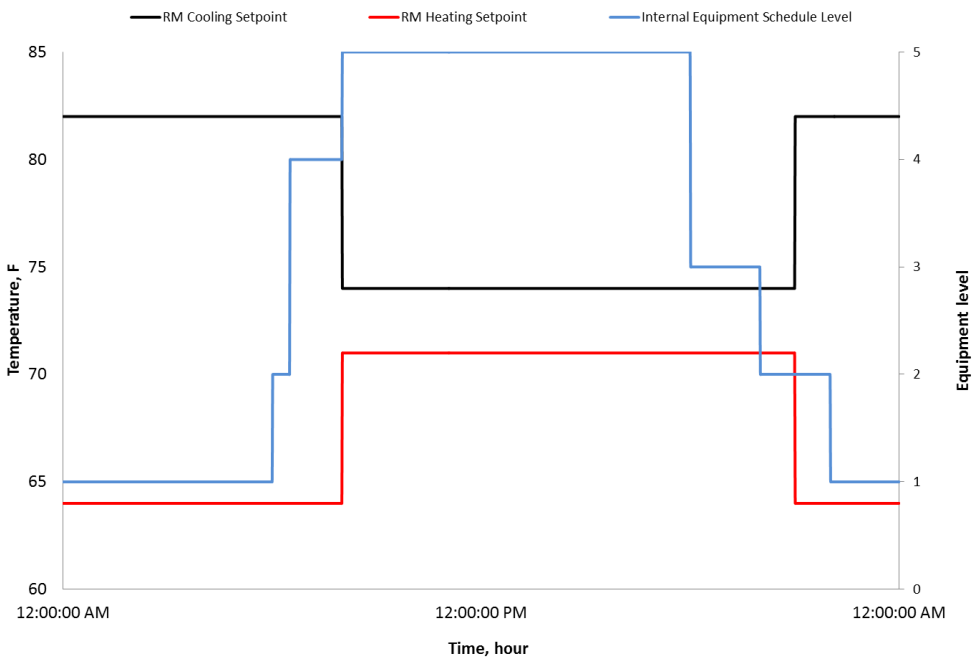


Figure 7-7. Real field system forecasting operation signals

7.4.3.1 Case I

In this case study, the training period is from 0:00 am 08/28/2015 to 9:00 pm 09/01/2015, and the forecasting period is from 9:00 pm 09/01/2015 to 9:00 pm 09/02/2015. The outside temperature range at the training period is 16.5 °C to 31.6 °C, while the temperature range at the forecasting period is 20.5 °C to 31.2 °C. On the other hand, there are sunny and cloudy days in the training period but the forecasting day is a sunny day. In the training period, the room temperature

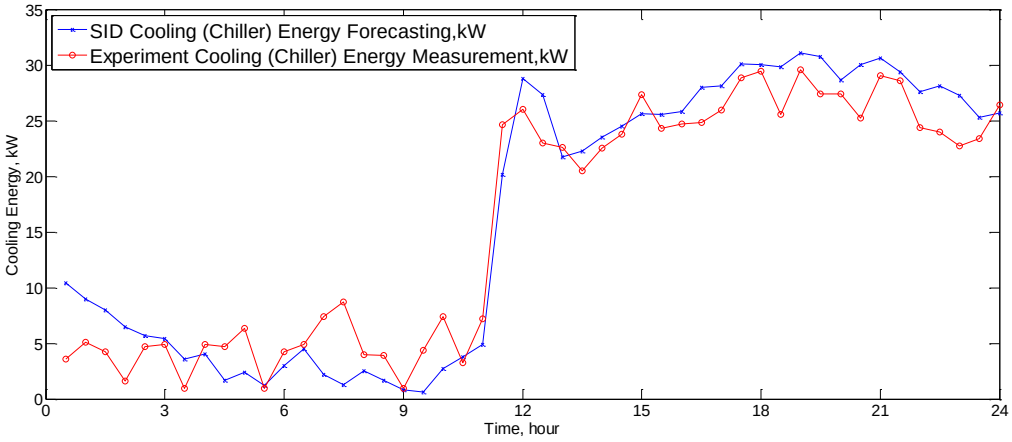
setpoints and equipment schedules are excited, however, common temperature setponts and equipment schedules (Figure 7-7) are used in the forecasting periods. The results are plotted in Figure 7-8a. In this figure, the general forecasting from the SID model is acceptable (Table 7-3), which during the unoccupied hours (0 - 9 am) the forecasting error is relative large.

7.4.3.2 Case II

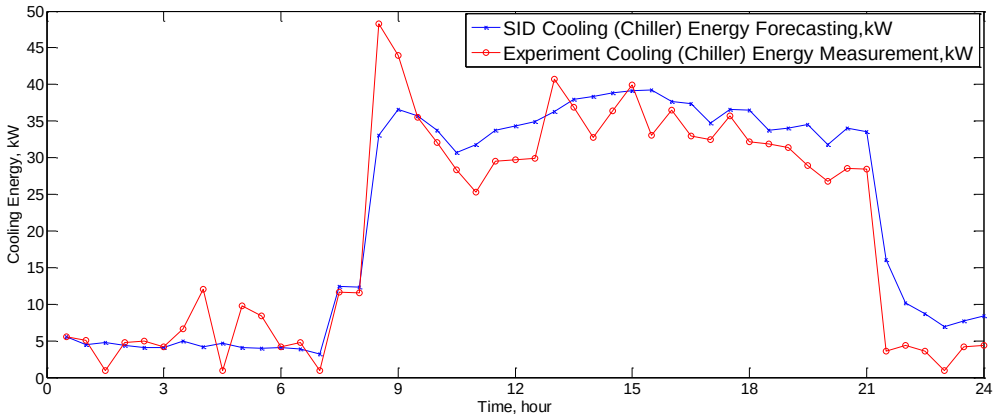
In this case study, the training period is from 0:00 am 08/31/2015 to 9:00 pm 09/01/2015, and 9:00 pm 09/02/2015 – 12:00 am 09/05/2015. The forecasting period is from 0:00 am 09/06/2015 – 12:00am 09/06/2015. The outside temperature range at the training period is 16.5 ℃ to 32.7 ℃, while the temperature range at the forecasting period is 20.5 ℃ to 31.2 ℃. The forecasting day is also a sunny day. Same operation signals (Figure 7-7) are used for this case. The results are plotted in Figure 7-8b. Similar to case I, the forecasting accuracy at unoccupied hours is larger than that in the occupied hours (Table 7-3)

7.4.3.3 Case III

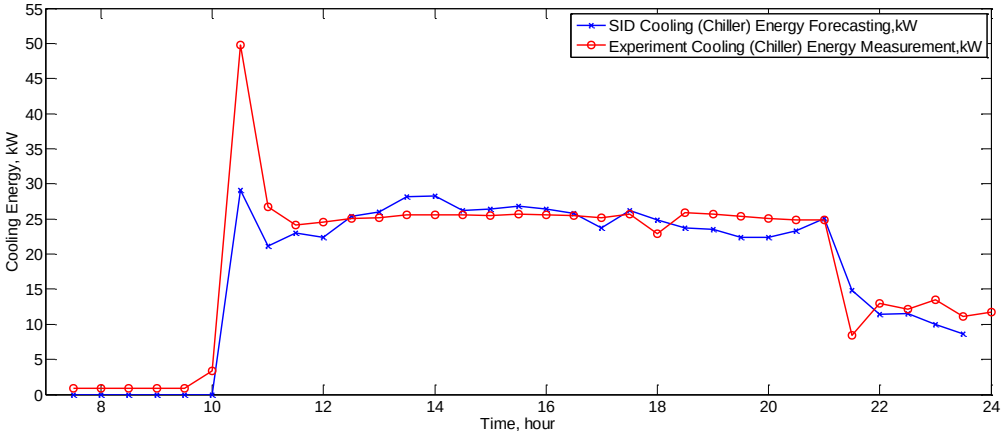
In this case study, the training data used in case II (section 7.4.3.2) is used in this scenario. The forecasting period is from 7:00 am 09/08/2015 – 12:00 am 09/06/2015. The outside temperature range is 18.6℃ to 26.3 ℃, which is also a sunny day. The operation signals in Figure 7-7 is also used in this study. The results are plotted in Figure 7-8c. Since there was a power outrage from 0:00 am to 7:00 am on 09/08/2015, all of the systems were off during this period. In Figure 7-8c, the SID model (blue line) is not able to capture this sudden increase of the cooling energy consumption (red line) at 9:00 am when the HVAC system was turned on from completely off.



a. Case I Forecasting results for 0901 to 0902



b. Case II forecasting results for 0906



a) Case III forecasting results for 0908

Figure 7-8. Real field SID energy forecasting results

Table 7-3 Real field cooling energy forecasting accuracy

Forecasting date	Overall			Occupied hours		
	R^2	RMSE, kW	NRMSE	R^2	RMSE, kW	NRMSE
0901 - 0902	0.92	3.1	10.0%	0.96	2.9	4.4%
0906	0.89	4.8	10.2%	0.94	3.1	6.8%
0908	0.86	4.6	9.0%	0.92	3.0	6.0%

In order to improve the model performance in unoccupied hours and system starting up & shutting down periods in case II and case III, a follow up experiment is then conducted to collect more building operation data in those operation conditions to enrich the model training data.

7.4.4 SID Model Refinement Experiment

As introduced, a follow up experiment is conducted from 09/15/2015 to 09/20/2015 to enrich the model training data in unoccupied hours and in starting up & shutting down periods. In this experiment, two day system nonlinearity test, two day system "on and off" test and three days normal test have been conducted. The nonlinearity test is majorly designed to test the nonlinearity between system output (cooling energy consumption) and solar radiation, because during the previous test, the radiometer was down. The system "on and off" test is to enrich the model training data under the system starting up and shutting down period to improve the model forecasting performance in these periods.

7.4.4.1 Demand response operation

In this operation, typical DR operation strategy recommended by [240] is applied in the test building (Table 7-4). Under this operation strategy, the measured zone temperature is plotted in Figure 7-9.

Table 7-4. DR operation temperature setpoint

Time	Cooling setpoint, ºC (ºF)	Heating setpoint, ºC (ºF)
0- 4 am	32 (89.6)	29 (84.2)
4 – 6 am	18 (64.4)	15 (59)
6 am – 6 pm	24 (75.2)	21 (69.8)
6 pm – 12 am	32 (89.6)	29 (84.2)

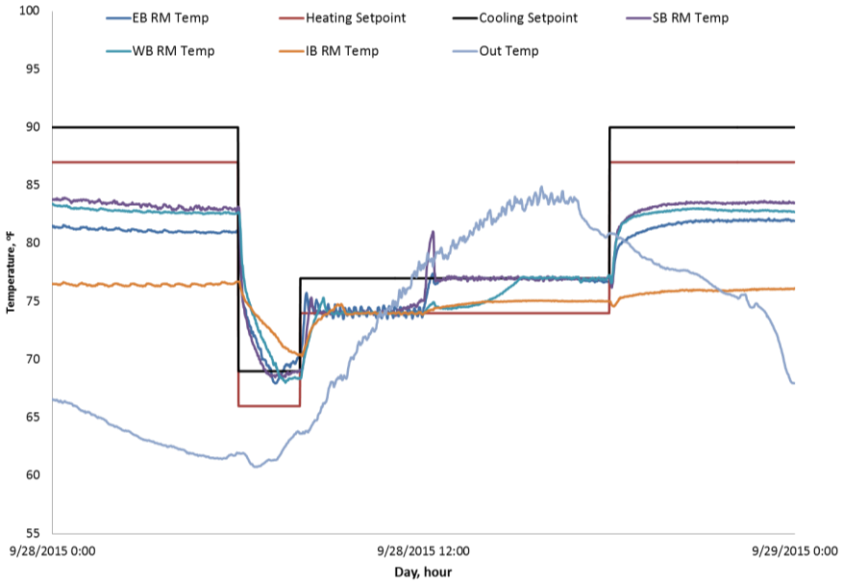


Figure 7-9. DR operation building temperature control

7.4.4.2 System on/off operation

As introduced in section 7.4.3, the HVAC system is switched on and off every 6 hours. The building temperature under this operation strategy is show in Figure 7-10.

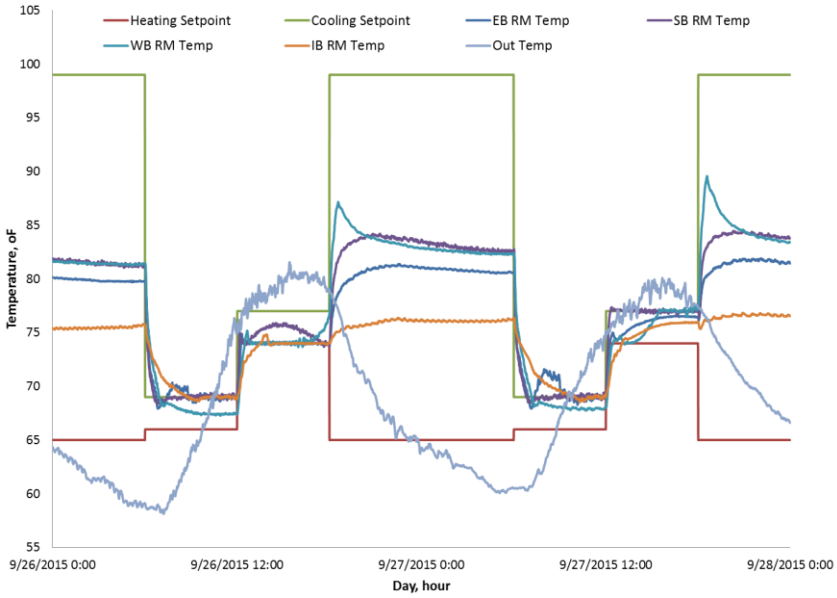
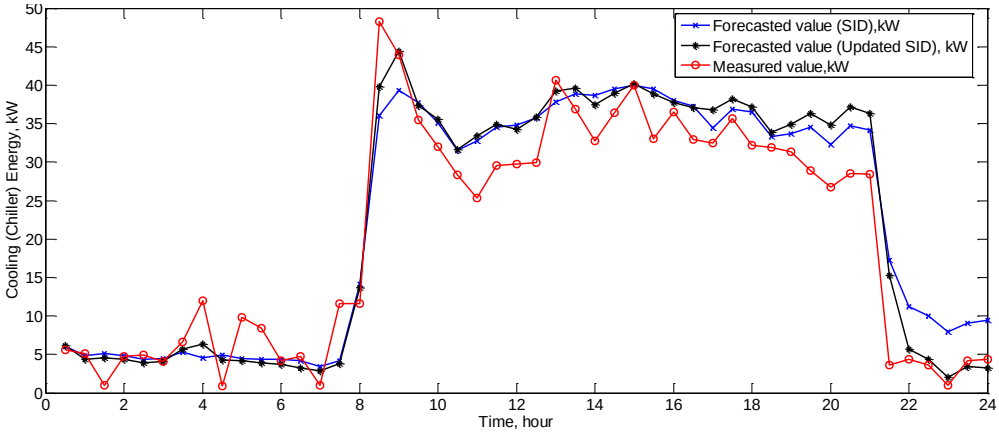


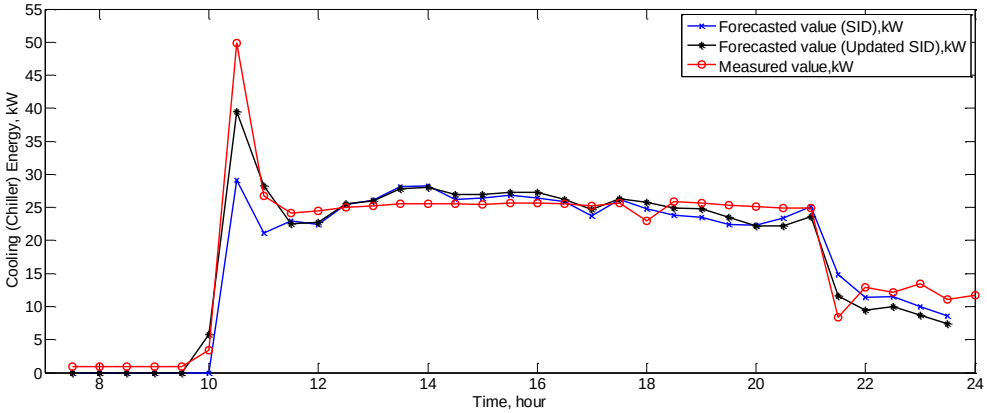
Figure 7-10. On-off operation building temperature control

7.4.4.3 SID model development and forecasting

The energy forecasting performance of the developed model as reported in section 7.4 is acceptable. But it is desired to improve its performance at the system starting up and shutting down periods. Therefore two days system on-off operation data and one day DR operation data is included into the training data set in this section. The energy forecasting results of the newly developed SID are plotted in Figure 7-11, where the updated SID model (black lines) achieved higher accuracy when forecasting the energy consumption at the system starting up and shutting down periods (9:00 am and 9:00 pm in Figure 7-11a, and 11:00 am and 9:00 pm in Figure 7-11b). When examining the energy forecasting results at the unoccupied hours, the fluctuation of the chiller energy consumption is found to be very large in the experimental data, which is believed to be caused by system faults. The SID model forecasts are more constant and closer to how the system supposes to perform



a) Case II forecasting results for 0906



b) Case III forecasting results for 0908

Figure 7-11. Real field updated SID energy forecasting results

Table 7-5. SID model energy forecasting accuracy

Case		Before updating			Updated		
		R^2	RMSE, kW	NRMSE	R^2	RMSE, kW	NRMSE
0906	Overall	0.85	4.8	10.2%	0.90	4.5	10.0%
	Occupied	0.96	3.9	8.4%	0.94	4.2	10.0%
0908	Overall	0.86	4.6	9.0%	0.94	2.7	9.0%
	Occupied	0.92	3.0	6.0%	0.95	2.0	3.3%

The energy forecasting result statistics are summarized in Table 7-5. In this table, the updated SID model improves the forecasting accuracy for all the cases, except for the occupied period of

0906. This is because the updated SID overestimated the cooling energy at the early evening from 18:00 to 21:00.

7.5 Conclusion

A whole building experiment is designed and conducted for 1) real building energy system characteristic test, and 2) real building system identification for energy forecasting model validation. The result shows that the developed methodology is able to evaluate the building energy system's characteristics and the system identification method is able to develop the cooling energy forecasting model with acceptable accuracy.

On the other hand, there are some limitations for the developed methodology to be applied in the real field:

1. In the system response time test, it is very hard to eliminate the outside disturbances such as temperature change and solar radiation disturbances.
2. In the system excitation period: except the temperature setpoint and equipment schedules, there are other control variables such as chilled water supply temperature and pressure, supply air temperature and pressure. All of these variables affect the cooling system operation and energy consumption. In this study, these other variables have not been considered. Therefore, in some of the test cases, the energy forecasting accuracy is not very good due to other variables' impacts. In the future, a systematic plan for modeling all major system input variables is needed.

8. CHAPTER 8 Summary and Future Work

8.1 Overall Achievement

The two main research objectives in this thesis are to develop a building cluster emulator for building cluster control and operation strategy development and to develop a novel methodology for high fidelity on-line building energy forecasting models development for real-time model based control and optimization.

1. Developed a building cluster emulator for “real-world-like” operation data generation and operation scheme assessment. Different physics based simulation models are interconnected and co-simulated through BCVTB. During the simulation, buildings, energy generation and storage systems are able to share information and exchange data. A proof-of-concept study is also conducted to validate the connection and data exchanging.
2. Developed and tested a novel on-line building energy forecasting methodology through proactive system identification and data fusion techniques. A frequency function based system identification methodology has been developed to forecast energy consumption. What’s more, a systematic approach for system characteristics test and model adaptation is also developed to improve the model robustness and extendibility.
3. Developed an on-line building cluster energy forecasting model using grey box modeling approach. Other than the on-line energy forecasting models for buildings, different models have also created and validated using grey box modeling approach for PV panel, battery and ice tank systems.
4. Applied and validated the above proposed methodologies under real world conditions. A comprehensive experiment has been conducted at Iowa energy center to verify the proposed system identification methodology for energy forecasting. In this experiment, a small-size

commercial building with VAV system is used for system characteristic test, system excitation, and system identification for energy forecasting model development and verification. The accuracy of the proposed system identification methodology is above 80% under the real conditions.

8.2 Limitations and Future Work

1. The emulator now has the interface place holder for the BACnet connection. Following work is necessary to verify the real connection and communicate with real building automation systems. Once the BACnet connection has been realized, the control strategies can developed in this emulator will be applied in the real building, and on the other hand, the real building operation data will be in-taken into the emulator for emulator simulation refinement and operation strategy assessment.
2. In this thesis, the building cluster emulator includes two different buildings, one ice tank energy storage system, and one PV and battery system. In the future work, this emulator can be expanded to include more buildings, and more transactive energy recourses for operation strategy development.
3. The on-line energy forecasting model developed in this thesis is only for cooling energy consumption (chiller energy). In the future, this model will need to be extended to the while building energy forecasting, including fan, boiler, lighting, equipment, etc.
4. The general system identification method for energy forecasting has only been applied and validated in three virtual buildings and a small size real building. In the future, this method will be tested in more buildings with different systems with different HVAC systems.

5. Apply the energy forecasting model in on-line model based control and operation optimization for energy consumption and cost saving and use the cluster emulator for the operation strategy assessment.

Reference

1. Parks, N., *Energy efficiency and the smart grid.* Environmental Science & Technology, 2009. **43**(9): p. 2999-3000.

2. Friedman, H., *Wiring the smart grid for energy savings: Integrating buildings to maximize investment*. 2009: Portland, Oregon: Portland Energy Conservation Inc. (PECI.).

3. DOE, U.S. *Buildings Energy Data Book, http://buildingsdatabook.eren.doe.gov/*. 2013 [cited 2015 01.10].

4. EISA, *Energy Independence and Security Act of 2007 Pub. L. No. 110-140, 121 Stat. 1492, 1783-84*. 2007.

5. EPBD, *The Directive 2010/31/EU of the European Parliament and of the Council of 19 May 2010 on the energy performance of buildings.* Official Journal of the European Union, 2010. **53**.

6. Kolokotsa, D., et al., *A roadmap towards intelligent net zero-and positive-energy buildings.* Solar Energy, 2011. **85**(12): p. 3067-3084.

7. National Energy Technology Laboratory (NETL), *Demand Dispatch–Intelligent Demand for a More Efficient Grid*, DOE/NETL-DE-FE0004001, Editor. August 2011: Morgantown, West Virginia.

8. Li, X. and J. Wen, *Review of building energy modeling for control and operation.* Renewable and Sustainable Energy Reviews, 2014. **37**: p. 517-537.

9. Office of Energy Efficiency and Renewable Energy, U.S.D.o.E. *EnergyPlus Energy Simulation Software: About EnergyPlus*. [cited 2014 March 03]; Available from: http://www.energyplus.gov.

10. TRNSYS. *Thermal energy system specialists*. Available from: http://www.trnsys.com.

11. Park, C., D.R. Clark, and G.E. Kelly. *An overview of HVACSIM+, a dynamic building/HVAC/control systems simulation program*. in *Proceedings of the 1st Annual Building Energy Simulation Conference, Seattle, WA*. 1985.

12. Wetter, M., *Modelica-based modelling and simulation to support research and development in building energy and control systems.* Journal of Building Performance Simulation, 2009. **2**(2): p. 143-161.

13. eQUEST. *The QUick Energy Simulation Tool*. Available from: http://www.doe2.com/equest/.

14. Crawley, D.B., et al., *Contrasting the capabilities of building energy performance simulation programs.* Building and Environment, 2008. **43(4):661–73.**

15. U.S. Department of Energy, E.E.a.R.E.O., Building Technology Program,Net-Zero Energy Commercial Building Initiative. *Commercial Building Benchmark Models. Available from: http://www1.eere.energy.gov/buildings/commercial_initiative/benchmark models.html*. 2009.

16. Asadi, E., et al., *A multi-objective optimization model for building retrofit strategies using TRNSYS simulations, GenOpt and MATLAB.* Building and Environment, 2012. **56**: p. 370-378.

17. Al-ajmi, F.F. and V.I. Hanby, *Simulation of energy consumption for Kuwaiti domestic buildings.* Energy and Buildings, 2008. **40**(6): p. 1101-1109.

18. Gonçalves, P., A.R. Gaspar, and M.G. da Silva, *Comparative energy and exergy performance of heating options in buildings under different climatic conditions.* Energy and Buildings, 2013. **61**: p. 288-297.

19. Khandelwal, A., P. Talukdar, and S. Jain, *Energy savings in a building using regenerative evaporative cooling.* Energy and Buildings, 2011. **43**(2–3): p. 581-591.

20. Shahrestani, M., R. Yao, and G.K. Cook, *Characterising the energy performance of centralised HVAC&R systems in the UK.* Energy and Buildings, 2013. **62**: p. 239-247.

21. Salvalai, G., J. Pfafferott, and M.M. Sesana, *Assessing energy and thermal comfort of different low-energy cooling concepts for non-residential buildings.* Energy Conversion and Management, 2013. **76**: p. 332-341.

22. Buonomano, A. and A. Palombo, *Building energy performance analysis by an in-house developed dynamic simulation code: An investigation for different case studies.* Applied Energy, 2014. **113**: p. 788-807.

23. Carpenter, J., et al., *Passive energy management through increased thermal capacitance.* Energy and Buildings, 2014. **75**: p. 465-471.

24. Buratti, C., et al., *Unsteady simulation of energy performance and thermal comfort in non-residential buildings.* Building and Environment, 2013. **59**: p. 482-491.

25. Li, S. and J. Wen, *Development and Validation of a Dynamic Air Handling Unit Model-Part I (RP 1312).* ASHRAE Transactions, 2010. **116(1)**: p. 45-56.

26. Wetter, M., W. Zuo, and T.S. Nouidui. *Modeling of heat transfer in rooms in the Modelica "Buildings" library*. in *Proceedings of Building Simulation 2011: 12th Conference of International Building Performance Simulation Association*. 2011. Sydney, 14-16 November.

27. Eisenhower, B., K. Gasljevic, and I. Mezic. *Control-oriented dynamic modeling and calibration of a campus theater using Modelica*. in *IBPSA-USA's SimBuild 2012 Conference The Future of Simulation for Buildings*. 2012.

28. Ansuini, R., et al. *Hybrid Modeling for Energy Saving in Subway Stations*. in *Proceedings of the 2012 Building Simulation and Optimization Conference, Loughborough*. 2012.

29. Tomažič, S., et al., *Indoor-environment simulator for control design purposes.* Building and Environment, 2013. **70**: p. 60-72.

30. Felgner, F., R. Merz, and L. Litz, *Modular modelling of thermal building behaviour using Modelica.* Mathematical and Computer Modelling of Dynamical Systems, 2006. **12**(1): p. 35-49.

31. Sodja, A. and B. Zupančič, *Modelling thermal processes in buildings using an object-oriented approach and Modelica.* Simulation Modelling Practice and Theory, 2009. **17**(6): p. 1143-1159.

32. Yin, R., P. Xu, and P. Shen, *Case study: Energy savings from solar window film in two commercial buildings in Shanghai.* Energy and Buildings, 2012. **45**: p. 132-140.

33. Ke, M.-T., C.-H. Yeh, and J.-T. Jian, *Analysis of building energy consumption parameters and energy savings measurement and verification by applying eQUEST software*. Energy and Buildings, 2013. **61**: p. 100-107.

34. Fuertes, G. and S. Schiavon, *Plug load energy analysis: The role of plug loads in LEED certification and energy modeling.* Energy and Buildings, 2014. **76**: p. 328-335.

35. Heiple, S. and D.J. Sailor, *Using building energy simulation and geospatial modeling techniques to determine high resolution building sector energy consumption profiles.* Energy and Buildings, 2008. **40**(8): p. 1426-1436.

36. Eltawil, M.A. and Z. Zhao, *Grid-connected photovoltaic power systems: Technical and potential problems—A review*. Renewable and Sustainable Energy Reviews, 2010. **14**(1): p. 112-129.

37. Ihm, P., M. Krarti, and G.P. Henze, *Development of a thermal energy storage model for EnergyPlus.* Energy and Buildings, 2004. **36**(8): p. 807-814.

38. Henze, G.P., M. Krarti, and M.J. Brandemuehl, *Guidelines for improved performance of ice storage systems.* Energy and Buildings, 2003. **35**(2): p. 111-127.

39. Candanedo, J.A., V.R. Dehkordi, and M. Stylianou, *Model-based predictive control of an ice storage device in a building cooling system.* Applied Energy, 2013. **111**: p. 1032-1045.

40. Sehar, F., S. Rahman, and M. Pipattanasomporn, *Impacts of ice storage on electrical energy consumptions in office buildings.* Energy and Buildings, 2012. **51**(0): p. 255-262.

41. Cameron, C.P., W.E. Boyson, and D.M. Riley. *Comparison of PV system performance-model predictions with measured PV system performance*. in *Photovoltaic Specialists Conference, 2008. PVSC'08. 33rd IEEE*. 2008. IEEE.

42. Saber, E.M., et al., *PV (photovoltaics) performance evaluation and simulation-based energy yield prediction for tropical buildings.* Energy, 2014. **71**: p. 588-595.

43. Hsieh, C.-M., et al., *Potential for installing photovoltaic systems on vertical and horizontal building surfaces in urban areas.* Solar Energy, 2013. **93**: p. 312-321.

44. Hachem, C., A. Athienitis, and P. Fazio, *Energy performance enhancement in multistory residential buildings.* Applied Energy, 2014. **116**: p. 9-19.

45. Stefanović, A., M. Bojić, and D. Gordić, *Achieving net zero energy cost house from old thermally non-insulated house using photovoltaic panels.* Energy and Buildings, 2014. **76**: p. 57-63.

46. Kalogirou, S.A., *Use of TRNSYS for modelling and simulation of a hybrid pv–thermal solar system for Cyprus.* Renewable Energy, 2001. **23**(2): p. 247-260.

47. Mondol, J.D., Y.G. Yohanis, and B. Norton, *Comparison of measured and predicted long term performance of grid a connected photovoltaic system.* Energy Conversion and Management, 2007. **48**(4): p. 1065-1080.

48. Choi, Y., et al., *PV Analyst: Coupling ArcGIS with TRNSYS to assess distributed photovoltaic potential in urban areas.* Solar Energy, 2011. **85**(11): p. 2924-2939.

49. Mondol, J.D., Y.G. Yohanis, and B. Norton, *Optimising the economic viability of grid-connected photovoltaic systems.* Applied Energy, 2009. **86**(7–8): p. 985-999.

50. Mondol, J.D., Y.G. Yohanis, and B. Norton, *The impact of array inclination and orientation on the performance of a grid-connected photovoltaic system.* Renewable Energy, 2007. **32**(1): p. 118-140.

51. Cao, S. and K. Sirén, *Impact of simulation time-resolution on the matching of PV production and household electric demand.* Applied Energy, 2014. **128**: p. 192-208.

52. Gallo, A., et al., *Analysis of Net Zero-energy Building in Spain. Integration of PV, Solar Domestic Hot Water and Air-conditioning Systems.* Energy Procedia, 2014. **48**: p. 828-836.

53. Reynders, G., T. Nuytten, and D. Saelens, *Potential of structural thermal mass for demand-side management in dwellings.* Building and Environment, 2013. **64**: p. 187-199.

54. May-Ostendorp, P., et al., *Model-predictive control of mixed-mode buildings with rule extraction.* Building and Environment, 2011. **46**(2): p. 428-437.

55. Corbin, C.D., G.P. Henze, and P. May-Ostendorp, *A model predictive control optimization environment for real-time commercial building application.* Journal of Building Performance Simulation 2013. **6**(3).

56. Yoon, J.H., R. Bladick, and A. Novoselac, *Demand response for residential buildings based on dynamic price of electricity.* Energy and Buildings, 2014. **80**: p. 531-541.

57. Wetter, M. *GenOpt®, Generic Optimization Program*. in *Seventh international IBPSA conference*. 2001.

58. Coffey, B., et al., *A software framework for model predictive control with GenOpt.* Energy and Buildings, 2010. **42**(7): p. 1084-1092.

59. Rackes, A. and M.S. Waring, *Using multiobjective optimizations to discover dynamic building ventilation strategies that can improve indoor air quality and reduce energy use.* Energy and Buildings, 2014. **75**: p. 272-280.

60. Seo, D., P. Ihm, and M. Krarti, *Development of an optimal daylighting controller.* Building and Environment, 2011. **46**(5): p. 1011-1022.

61. Cole, W.J., E.T. Hale, and T.F. Edgar. *Building energy model reduction for model predictive control using OpenStudio*. in *2013 American Control Conference*. 2013. Washington, D.C.

62. Catalina, T., J. Virgone, and E. Blanco, *Development and validation of regression models to predict monthly heating demand for residential buildings.* Energy and Buildings, 2008. **40**(10): p. 1825-1832.

63. Pagliarini, G. and S. Rainieri, *Restoration of the building hourly space heating and cooling loads from the monthly energy consumption.* Energy and Buildings, 2012. **49**: p. 348-355.

64. Wan, K.K.W., et al., *Impact of modelled global solar radiation on simulated building heating and cooling loads.* Energy Conversion and Management, 2009. **50**(3): p. 662-667.

65. Wan, K.K.W., et al., *Future trends of building heating and cooling loads and energy consumption in different climates.* Building and Environment, 2011. **46**(1): p. 223-234.

66. Sullivan, R., et al., *Commercial building energy performance analysis using multiple regression.* ASHRAE Transactions, 1985. **91**: p. 337.

67. Markus, T.A., et al., *The influence of climate on housing: A simple technique for the assessment of dynamic energy behaviour.* Energy and Buildings, 1984. **7**(3): p. 243-259.

68. Lam, J.C., R.Y. Chan, and D.H. Li, *Simple regression models for fully air-conditioned public sector office buildings in subtropical climates.* Architectural Science Review, 2002. **45**(4): p. 361-369.

69. Korolija, I., et al., *Regression models for predicting UK office building energy consumption from heating and cooling demands.* Energy and Buildings, 2013. **59**: p. 214-227.

70. Masuda, H. and D.E. Claridge, *Statistical modeling of the building energy balance variable for screening of metered energy use in large commercial buildings.* Energy and Buildings, 2014. **77**: p. 292-303.

71. Moroşan, P.-D., et al., *Building temperature regulation using a distributed model predictive control.* Energy and Buildings, 2010. **42**(9): p. 1445-1452.

72. Yun, K., et al., *Building hourly thermal load prediction using an indexed ARX model.* Energy and Buildings, 2012. **54**: p. 225-233.

73. Wallace, M., et al., *Energy efficient model predictive building temperature control.* Chemical Engineering Science, 2012. **69**(1): p. 45-58.

74. Guo, Y., et al., *Hourly cooling load forecasting using time-indexed ARX models with two-stage weighted least squares regression.* Energy Conversion and Management, 2014. **80**(0): p. 46-53.

75. Ríos-Moreno, G.J., et al., *Modelling temperature in intelligent buildings by means of autoregressive models.* Automation in Construction, 2007. **16**(5): p. 713-722.

76. Ma, J., et al., *Demand reduction in building energy systems based on economic model predictive control.* Chemical Engineering Science, 2012. **67**(1): p. 92-100.

77. Kalogirou, S.A. and M. Bojic, *Artificial neural networks for the prediction of the energy consumption of a passive solar building.* Energy, 2000. **25**(5): p. 479-491.

78. Kalogirou, S.A., *Applications of artificial neural-networks for energy systems.* Applied Energy, 2000. **67**(1–2): p. 17-35.

79. Zhao, H.-x. and F. Magoulès, *A review on the prediction of building energy consumption.* Renewable and Sustainable Energy Reviews, 2012. **16**(6): p. 3586-3592.

80. Aydinalp, M., V. Ismet Ugursal, and A.S. Fung, *Modeling of the appliance, lighting, and space-cooling energy consumptions in the residential sector using neural networks.* Applied Energy, 2002. **71**(2): p. 87-110.

81. Aydinalp, M., V. Ismet Ugursal, and A.S. Fung, *Modeling of the space and domestic hot-water heating energy-consumption in the residential sector using neural networks.* Applied Energy, 2004. **79**(2): p. 159-178.

82. Yang, J., H. Rivard, and R. Zmeureanu, *On-line building energy prediction using adaptive artificial neural networks.* Energy and Buildings, 2005. **37**(12): p. 1250-1259.

83. Kwok, S.S.K. and E.W.M. Lee, *A study of the importance of occupancy to building cooling load in prediction by intelligent approach.* Energy Conversion and Management, 2011. **52**(7): p. 2555-2564.

84. Moon, J.W. and J.-J. Kim, *ANN-based thermal control models for residential buildings.* Building and Environment, 2010. **45**(7): p. 1612-1625.

85. Yokoyama, R., T. Wakui, and R. Satake, *Prediction of energy demands using neural network with model identification by global optimization.* Energy Conversion and Management, 2009. **50**(2): p. 319-327.

86. Dong, B., C. Cao, and S.E. Lee, *Applying support vector machines to predict building energy consumption in tropical region.* Energy and Buildings, 2005. **37**(5): p. 545-553.

87. Li, Q., et al., *Applying support vector machine to predict hourly cooling load in the building.* Applied Energy, 2009. **86**(10): p. 2249-2256.

88. Zhao, H.X. and F. Magoulès, *Parallel support vector machines applied to the prediction of multiple buildings energy consumption.* Journal of Algorithms & Computational Technology, 2010. **4**(2): p. 231-249.

89. Lai, F., F. Magoules, and F. Lherminier, *Vapnik's learning theory applied to energy consumption forecasts in residential buildings.* International Journal of Computer Mathematics, 2008. **85**(10): p. 1563-1588.

90. Hou, Z. and Z. Lian. *An application of support vector machines in cooling load prediction*. in *Intelligent Systems and Applications, 2009. ISA 2009. International Workshop on*. 2009. IEEE.

91. Bacher, P., H. Madsen, and H.A. Nielsen, *Online short-term solar power forecasting.* Solar Energy, 2009. **83**(10): p. 1772-1783.

92. Giraud, F. and Z.M. Salameh, *Analysis of the effects of a passing cloud on a grid-interactive photovoltaic system with battery storage using neural networks.* Energy Conversion, IEEE Transactions on, 1999. **14**(4): p. 1572-1577.

93. Mellit, A., M. Benghanem, and S.A. Kalogirou, *Modeling and simulation of a stand-alone photovoltaic system using an adaptive artificial neural network: Proposition for a new sizing procedure.* Renewable Energy, 2007. **32**(2): p. 285-313.

94. Kalogirou, S.A., E. Mathioulakis, and V. Belessiotis, *Artificial neural networks for the performance prediction of large solar systems.* Renewable Energy, 2014. **63**: p. 90-97.

95. El-Sawi, A., F. Haghighat, and H. Akbari, *Assessing long-term performance of centralized thermal energy storage system.* Applied Thermal Engineering, 2014. **62**(2): p. 313-321.

96. Giraud, F. and Z.M. Salameh, *Combined effects of passing clouds and wind gusts on an interactive wind-PV system with battery storage using neural networks.* Electric Power Components and Systems, 2007. **35**(7): p. 823-836.

97. Zeng, J. and W. Qiao, *Short-term solar power prediction using a support vector machine.* Renewable Energy, 2013. **52**: p. 118-127.

98. Bouzerdoum, M., A. Mellit, and A. Massi Pavan, *A hybrid model (SARIMA–SVM) for short-term power forecasting of a small-scale grid-connected photovoltaic plant.* Solar Energy, 2013. **98, Part C**: p. 226-235.

99. Álvarez Antón, J.C., et al., *Battery state-of-charge estimator using the SVM technique.* Applied Mathematical Modelling, 2013. **37**(9): p. 6244-6253.

100. Avci, M., et al., *Model predictive HVAC load control in buildings using real-time electricity pricing.* Energy and Buildings, 2013. **60**: p. 199-209.

101. Cole, W.J., et al., *Reduced-order residential home modeling for model predictive control.* Energy and Buildings, 2014. **74**: p. 69-77.

102. Shi, J., J. Guo, and S. Zheng, *Evaluation of hybrid forecasting approaches for wind speed and power generation time series.* Renewable and Sustainable Energy Reviews, 2012. **16**(5): p. 3471-3480.

103. Braun, J.E., *Reducing Energy Costs and Peak Electrical Demand through Optimal Control of Building Thermal Storage.* ASHRAE Transactions, 1990. **96(2)**: p. 876-888.

104. Moon, J.W., *Performance of ANN-based predictive and adaptive thermal-control methods for disturbances in and around residential buildings.* Building and Environment, 2012. **48**: p. 15-26.

105. Hong, C.-M., F.-S. Cheng, and C.-H. Chen, *Optimal control for variable-speed wind generation systems using General Regression Neural Network.* International Journal of Electrical Power & Energy Systems, 2014. **60**: p. 14-23.

106. Xi, X.-C., A.-N. Poo, and S.-K. Chou, *Support vector regression model predictive control on a HVAC plant.* Control Engineering Practice, 2007. **15**(8): p. 897-908.

107. Eisenhower, B., et al., *A methodology for meta-model based optimization in building energy models.* Energy and Buildings, 2012. **47**: p. 292-301.

108. Tutkun, N., *Minimization of operational cost for an off-grid renewable hybrid system to generate electricity in residential buildings through the SVM and the BCGA methods.* Energy and Buildings, 2014. **76**: p. 470-475.

109. Yushan, L., et al., *Control System Design of Battery-Assisted Quasi-Z-Source Inverter for Grid-Tie Photovoltaic Power Generation.* Sustainable Energy, IEEE Transactions on, 2013. **4**(4): p. 994-1001.

110. Braun, J.E. and N. Chaturvedi, *An inverse gray-box model for transient building load prediction.* HVAC&R Research, 2002. **8**(1): p. 73-99.

111. Henze, G.P., C. Felsmann, and G. Knabe, *Evaluation of optimal control for active and passive building thermal storage.* International Journal of Thermal Sciences, 2004. **43**(2): p. 173-183.

112. Wang, S. and X. Xu, *Simplified building model for transient thermal performance estimation using GA-based parameter identification.* International Journal of Thermal Sciences, 2006. **45**(4): p. 419-432.

113. Zhou, Q., et al., *A grey‐box model of next‐day building thermal load prediction for energy‐efficient control.* International Journal of Energy Research, 2008. **32**(15): p. 1418-1431.

114. Bueno, B., et al., *A resistance-capacitance network model for the analysis of the interactions between the energy performance of buildings and the urban climate.* Building and Environment, 2012. **54**: p. 116-125.

115. Hazyuk, I., C. Ghiaus, and D. Penhouet, *Optimal temperature control of intermittently heated buildings using Model Predictive Control: Part I – Building modeling.* Building and Environment, 2012. **51**: p. 379-387.

116. Gouda, M.M., S. Danaher, and C.P. Underwood, *Building thermal model reduction using nonlinear constrained optimization.* Building and Environment, 2002. **37**(12): p. 1255-1265.

117. Dong, B., et al., *Development and calibration of an online energy model for campus buildings.* Energy and Buildings, 2014. **76**: p. 316-327.

118. Lee, K.-h. and J.E. Braun, *Model-based demand-limiting control of building thermal mass.* Building and Environment, 2008. **43**(10): p. 1633-1646.

119. Deng, K., et al., *Structure-preserving model reduction of nonlinear building thermal models.* Automatica, 2014. **50**(4): p. 1188-1195.

120. Wen, J. and T.F. Smith, *Development and validation of online models with parameter estimation for a building zone with VAV system.* Energy and Buildings, 2007. **39**(1): p. 13-22.

121. Wang, Y.-W., et al., *A simplified modeling of cooling coils for control and optimization of HVAC systems.* Energy Conversion and Management, 2004. **45**(18–19): p. 2915-2930.

122. Platt, G., et al., *Adaptive HVAC zone modeling for sustainable buildings.* Energy and Buildings, 2010. **42**(4): p. 412-421.

123. Jin, G.-Y., et al., *A simplified modeling of mechanical cooling tower for control and optimization of HVAC systems.* Energy Conversion and Management, 2007. **48**(2): p. 355-365.

124. Beghi, A. and L. Cecchinato, *Modelling and adaptive control of small capacity chillers for HVAC applications.* Applied Thermal Engineering, 2011. **31**(6–7): p. 1125-1134.

125. Henze, G.P., et al., *Impact of forecasting accuracy on predictive optimal control of active and passive building thermal storage inventory.* HVAC&R Research, 2004. **10**(2): p. 153-178.

126. Lee, W.-S., Y.-T. Chen, and T.-H. Wu, *Optimization for ice-storage air-conditioning system using particle swarm algorithm.* Applied Energy, 2009. **86**(9): p. 1589-1595.

127. Chen, H.-J., D.W.P. Wang, and S.-L. Chen, *Optimization of an ice-storage air conditioning system using dynamic programming method.* Applied Thermal Engineering, 2005. **25**(2–3): p. 461-472.

128. Rismanchi, B., et al., *Modeling and simulation to determine the potential energy savings by implementing cold thermal energy storage system in office buildings.* Energy Conversion and Management, 2013. **75**(0): p. 152-161.

129. Zhanbo, X., et al., *Performance Analysis and Comparison on Energy Storage Devices for Smart Building Energy Management.* Smart Grid, IEEE Transactions on, 2012. **3**(4): p. 2136-2147.

130. Lu, L. and H. Yang, *A study on simulations of the power output and practical models for building integrated photovoltaic systems.* Journal of solar energy engineering, 2004. **126**(3): p. 929-935.

131. Karatepe, E., M. Boztepe, and M. Çolak, *Development of a suitable model for characterizing photovoltaic arrays with shaded solar cells.* Solar Energy, 2007. **81**(8): p. 977-992.

132. Henze, G.P. and R.H. Dodier, *Adaptive optimal control of a grid-independent photovoltaic system.* Journal of solar energy engineering, 2003. **125**(1): p. 34-42.

133. Ban-Weiss, G., et al., *Electricity production and cooling energy savings from installation of a building-integrated photovoltaic roof on an office building.* Energy and Buildings, 2013. **56**: p. 210-220.

134. Zahedi, A., *Maximizing solar PV energy penetration using energy storage technology.* Renewable and Sustainable Energy Reviews, 2011. **15**(1): p. 866-870.

135. Oldewurtel, F., et al., *Use of model predictive control and weather forecasts for energy efficient building climate control.* Energy and Buildings, 2012. **45**: p. 15-27.

136. Lee, K.-h. and J.E. Braun. *Reducing peak cooling loads through model-based control of zone temperature setpoints*. in *American Control Conference, 2007. ACC'07*. 2007. IEEE.

137. Široký, J., et al., *Experimental analysis of model predictive control for an energy efficient building heating system.* Applied Energy, 2011. **88**(9): p. 3079-3087.

138. Ma, Y., et al., *Model predictive control for the operation of building cooling systems.* Control Systems Technology, IEEE Transactions on, 2012. **20**(3): p. 796-803.

139. Hu, J. and P. Karava, *Model predictive control strategies for buildings with mixed-mode cooling.* Building and Environment, 2014. **71**: p. 233-244.

140. Freire, R.Z., G.H.C. Oliveira, and N. Mendes, *Predictive controllers for thermal comfort optimization and energy savings.* Energy and Buildings, 2008. **40**(7): p. 1353-1365.

141. Welch, R. and G. Venayagamoorthy. *Comparison of two optimal control strategies for a grid independent photovoltaic system*. in *Industry Applications Conference, 2006. 41st IAS Annual Meeting. Conference Record of the 2006 IEEE*. 2006. IEEE.

142. Zervas, P., et al., *Model-based optimal control of a hybrid power generation system consisting of photovoltaic arrays and fuel cells.* Journal of power sources, 2008. **181**(2): p. 327-338.

143. Vatau, D., et al., *PV systems modelling and optimal control.* Energy Conversion and Management, 2014. **84**: p. 448-456.

144. Henze, G.P., *Impact of real-time pricing rate uncertainty on the annual performance of cool storage systems.* Energy and Buildings, 2003. **35**(3): p. 313-325.

145. Ma, Y., et al. *Model predictive control of thermal energy storage in building cooling systems*. in *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*. 2009. IEEE.

146. Hu, M., J.D. Weir, and T. Wu, *Decentralized operation strategies for an integrated building energy system using a memetic algorithm.* European Journal of Operational Research, 2012. **217**(1): p. 185-197.

147. Henze, G.P., et al., *Experimental analysis of model-based predictive optimal control for active and passive building thermal storage inventory.* HVAC&R Research, 2005. **11**(2): p. 189-213.

148. Vetterli, J. and M. Benz, *Cost-optimal design of an ice-storage cooling system using mixed-integer linear programming techniques under various electricity tariff schemes.* Energy and Buildings, 2012. **49**(0): p. 226-234.

149. Lennart, L., *System identification: theory for the user.* PTR Prentice Hall, Upper Saddle River, NJ, 1999.

150. Rivera, D.E., et al., *Constrained multisine input signals for plant-friendly identification of chemical process systems.* Journal of Process Control, 2009. **19**(4): p. 623-635.

151. Subudhi, B. and D. Jena, *A differential evolution based neural network approach to nonlinear system identification.* Applied Soft Computing, 2011. **11**(1): p. 861-871.

152. Schafroth, D., et al., *Modeling, system identification and robust control of a coaxial micro helicopter.* Control Engineering Practice, 2010. **18**(7): p. 700-711.

153. Simani, S. and R.J. Patton, *Fault diagnosis of an industrial gas turbine prototype using a system identification approach.* Control Engineering Practice, 2008. **16**(7): p. 769-786.

154. Alfi, A. and H. Modares, *System identification and control using adaptive particle swarm optimization.* Applied Mathematical Modelling, 2011. **35**(3): p. 1210-1221.

155. Privara, S., et al. *Modeling and identification of a large multi-zone office building*. in *Control Applications (CCA), 2011 IEEE International Conference on*. 2011. IEEE.

156. Privara, S., et al. *Predictive control oriented subspace identification based on building energy simulation tools*. in *Control & Automation (MED), 2012 20th Mediterranean Conference on*. 2012. IEEE.

157. Zacekova, E., S. Privara, and Z. Vana. *Model predictive control relevant identification using partial least squares for building modeling*. in *Australian Control Conference (AUCC), 2011*. 2011. IEEE.

158. Braun, M., et al. *Multi-level pseudo-random signal design and "model-on-demand" estimation applied to nonlinear identification of a RTP wafer reactor*. in *American Control Conference, 1999. Proceedings of the 1999*. 1999. IEEE.

159. Braun, M., R. Ortiz-Mojica, and D. Rivera, *Application of minimum crest factor multisinusoidal signals for "plant-friendly" identification of nonlinear process systems.* Control Engineering Practice, 2002. **10**(3): p. 301-313.

160. Agbi, C., Z. Song, and B. Krogh. *Parameter identifiability for multi-zone building models*. in *Decision and Control (CDC), 2012 IEEE 51st Annual Conference on*. 2012. IEEE.

161. Castanedo, F., *A review of data fusion techniques.* The Scientific World Journal, 2013. **2013**.

162. O'Neill, Z., S. Narayanan, and R. Brahme, *Model-based thermal load estimation in buildings.* Proceedings of SimBuild, 2010.

163. Fux, S.F., et al., *EKF based self-adaptive thermal model for a passive house.* Energy and Buildings, 2014. **68, Part C**: p. 811-817.

164. Bernal, W., et al. *MLE+: a tool for integrated design and deployment of energy efficient building controls*. in *Proceedings of the Fourth ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*. 2012. ACM.

165. Wetter, M., *Co-simulation of building energy and control systems with the Building Controls Virtual Test Bed.* Journal of Building Performance Simulation, 2011. **4**(3): p. 185-203.

166. Pang, X., et al., *A framework for simulation-based real-time whole building performance assessment.* Building and Environment, 2012. **54**: p. 100-108.

167. Nouvel, R. and F. Alessi. *A novel personalized thermal comfort control, responding to user sensation feedbacks*. in *Building Simulation*. 2012. Springer.

168. MathWorks, I., *MATLAB and Statistics Toolbox Release 2012b*. 2012: Natick, Massachusetts, United States.

169. Liam Hendricken, et al., *Capital Costs and Energy Savings Achieved by Energy Conservation Measures for Office Buildings in the Greater Philadephia Region*, in *SimBuild* 2012: Madison, WI.

170. Fry, B.A., *Simulation of Grid-Tied Building Integrated Photovoltaic Systems.* Master's Thesis, University of Wisconsin-Madison, 1998.

171. DOE, U., *EnergyPlus Engineering Reference.* Input and Output Reference, 2013: p. 6.

172. Dong, B., Z. O'Neill, and Z. Li, *A BIM-enabled information infrastructure for building energy Fault Detection and Diagnostics.* Automation in Construction, 2014. **44**: p. 197-211.

173. Prívara, S., et al., *Building modeling: Selection of the most appropriate model for predictive control.* Energy and Buildings, 2012.

174. Ljung, L., *System identification*. 1999: Wiley Online Library.

175. Rivera, D.E., X. Chen, and D.S. Bayard. *Experimental design for robust process control using schroeder-phased input signals*. in *American Control Conference, 1993*. 1993. IEEE.

176. Devore, J.L., *Probability and Statistics for Engineering and the Sciences*, ed. 6. 2004: Brooks/Cole Publishing Co.

177. Deru, M., et al., *US Department of Energy commercial reference building models of the national building stock.* 2011.

178. Haber, R., *Nonlinear System Identification: Input-Output Modeling Approach, Vol. 2: Nonlinear System Structure Identification*, ed. 1. 1999: Springer.

179. Van Doren, J.F., et al. *Identifiability: from qualitative analysis to model structure approximation*. in *Proceedings of the 15th IFAC Symposium on System Identification, Saint-Malo, France*. 2009.

180. Juang, J.N., *Applied system identification*. 1994, Englewood Cliffs, New Jersey: Prentice Hall.

181. Li, X. and J. Wen, *Building energy consumption on-line forecasting using physics based system identification.* Energy and Buildings, 2014. **82**: p. 1-12.

182. Li, X., J. Wen, and T. Wu, *Net-zero Energy Impact Building Clusters Emulator for Operation Strategy Development*, in *2014 ASHRAE Annual Conference* 2014: Seattle, WA.

183. ASHRAE, *ASHRAE Fundamentals Handbook.* American Society of Heating, Refrigerating and Air Conditioning Engineers, Atlanta, 2001. **111**.

184. Van Overschee, P. and B. De Moor, *N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems.* Automatica, 1994. **30**(1): p. 75-93.

185. Braun, J.E., *A near-optimal control strategy for cool storage systems with dynamic electric rates (RP-1252).* HVAC&R Research, 2007. **13**(4): p. 557-580.

186. Chang, C.-C. and C.-J. Lin, *LIBSVM: a library for support vector machines.* ACM Transactions on Intelligent Systems and Technology (TIST), 2011. **2**(3): p. 27.

187. Ben-Nakhi, A.E. and M.A. Mahmoud, *Energy conservation in buildings through efficient A/C control using neural networks.* Applied Energy, 2002. **73**(1): p. 5-23.

188. Beale, M.H., M.T. Hagan, and H.B. Demuth, *Neural Network Toolbox 7.* User's Guide, MathWorks, 2010.

189. Van Overschee, P. and B. De Moor, *N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems.* Automatica, 1994. **30**(1): p. 75-93.

190. Prívara, S., VáHa, Z., Gyalistras, D., Cigler, J., Sagerschnig, C., Morari, M., & Ferkl, L., *Modeling and identification of a large multi-zone office building*, in *Control Applications (CCA), 2011 IEEE International Conference*. 2011.

191. Halvgaard, R., Poulsen, N. K., Madsen, H., & Jørgensen, J. B. , *Economic model predictive control for building climate control in a smart grid.* , in *In Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES*. 2012 IEEE. p. 1-6.

192. West, J.D. and J.E. Braun, *Modeling partial charging and discharging of areaconstrained ice storage tanks.* HVAC&R Research, 1999. **5**: p. 20.

193. Henze, G.P., M. Krarti, and M.J. Brandemuehl, *A simulation environment for the analysis of ice storage controls.* HVAC&R Research, 1997. **3**(2): p. 128-148.

194. Hajiah, A. and M. Krarti, *Optimal control of building storage systems using both ice storage and thermal mass–Part I: Simulation environment.* Energy Conversion and Management, 2012.

195. Hajiah, A. and M. Krarti, *Optimal controls of building storage systems using both ice storage and thermal mass–Part II: Parametric analysis.* Energy Conversion and Management, 2012.

196. Chen, H.J., D.W.P. Wang, and S.L. Chen, *Optimization of an ice-storage air conditioning system using dynamic programming method.* Applied thermal engineering, 2005. **25**(2): p. 461-472.

197. Massie, D.D., *Optimization of a building's cooling plant for operating cost and energy use*. International Journal of Thermal Sciences, 2002. **41**(12): p. 1121-1129.

198. Martín, L., et al., *Prediction of global solar irradiance based on time series analysis: Application to solar thermal power plants energy production planning*. Solar Energy, 2010. **84**(10): p. 1772-1781.

199. Yoshida, H., et al., *Modeling of Weather Data by Time Series Analysis for Air-conditioning Load Calculations*. 1992: American Society of Heating, Refrigerating and Air-conditioning Engineers.

200. Reikard, G., *Predicting solar radiation at high resolutions: A comparison of time series forecasts.* Solar Energy, 2009. **83**(3): p. 342-349.

201. Paoli, C., et al., *Forecasting of preprocessed daily solar radiation time series using neural networks.* Solar Energy, 2010. **84**(12): p. 2146-2160.

202. Voyant, C., et al., *Optimization of an artificial neural network dedicated to the multivariate forecasting of daily global radiation.* Energy, 2011. **36**(1): p. 348-359.

203. Kaushika, N.D., R.K. Tomar, and S.C. Kaushik, *Artificial neural network model based on interrelationship of direct, diffuse and global solar radiations.* Solar Energy, 2014. **103**(0): p. 327-342.

204. Liu, B.Y.H. and R.C. Jordan, *The interrelationship and characteristic distribution of direct, diffuse and total solar radiation.* Solar Energy, 1960. **4**(3): p. 1-19.

205. Loutzenhiser, P., et al., *Empirical validation of models to compute solar irradiance on inclined surfaces for building energy simulation.* Solar Energy, 2007. **81**(2): p. 254-267.

206. Coskun, C., Z. Oktay, and I. Dincer, *Estimation of monthly solar radiation distribution for solar energy system analysis.* Energy, 2011. **36**(2): p. 1319-1323.

207. Ulgen, K. and A. Hepbasli, *Diffuse solar radiation estimation models for Turkey's big cities.* Energy Conversion and Management, 2009. **50**(1): p. 149-156.

208. Yang, D., et al., *Solar irradiance forecasting using spatial-temporal covariance structures and time-forward kriging.* Renewable Energy, 2013. **60**(0): p. 235-245.

209. Boland, J., J. Huang, and B. Ridley, *Decomposing global solar radiation into its direct and diffuse components.* Renewable and Sustainable Energy Reviews, 2013. **28**(0): p. 749-756.

210. Perez, R., et al., *A new operational model for satellite-derived irradiances: description and validation.* Solar Energy, 2002. **73**(5): p. 307-317.

211. Ridley, B., J. Boland, and P. Lauret, *Modelling of diffuse solar fraction with multiple predictors.* Renewable Energy, 2010. **35**(2): p. 478-483.

212. Olmo, F.J., et al., *Prediction of global irradiance on inclined surfaces from horizontal global irradiance.* Energy, 1999. **24**(8): p. 689-704.

213. Redweik, P., C. Catita, and M. Brito, *Solar energy potential on roofs and facades in an urban landscape.* Solar Energy, 2013. **97**(0): p. 332-341.

214. Bakirci, K., *Models of solar radiation with hours of bright sunshine: A review.* Renewable and Sustainable Energy Reviews, 2009. **13**(9): p. 2580-2588.

215. Khalil, S.A. and A.M. Shaffie, *A comparative study of total, direct and diffuse solar irradiance by using different models on horizontal and inclined surfaces for Cairo, Egypt.* Renewable and Sustainable Energy Reviews, 2013. **27**: p. 853-863.

216. Badescu, V., *A new kind of cloudy sky model to compute instantaneous values of diffuse and global solar irradiance.* Theoretical and Applied Climatology, 2002. **72**(1-2): p. 127-136.

217. Tian, Y., et al., *Estimating solar radiation on slopes of arbitrary aspect.* Agricultural and Forest Meteorology, 2001. **109**(1): p. 67-74.

218. Skartveit, A. and J. Asle Olseth, *Modelling slope irradiance at high latitudes.* Solar Energy, 1986. **36**(4): p. 333-344.

219. Steven, M. and M.H. Unsworth, *The angular distribution and interception of diffuse solar radiation below overcast skies.* Quarterly Journal of the Royal Meteorological Society, 1980. **106**(447): p. 57-61.

220. Perez, R., et al., *Modeling daylight availability and irradiance components from direct and global irradiance.* Solar Energy, 1990. **44**(5): p. 271-289.

221. King, D., W. Boyson, and J. Kratochvill, *Photovoltaic Array Performance Model*. 2004, Sandia National Laboratories.

222. Davis, M.W., A.H. Fanney, and B.P. Dougherty, *Measured versus predicted performance of building integrated photovoltaics.* TRANSACTIONS-AMERICAN SOCIETY OF

MECHANICAL ENGINEERS JOURNAL OF SOLAR ENERGY ENGINEERING, 2003. **125**(1): p. 21-27.

223. Jones, A. and C. Underwood, *A thermal model for photovoltaic systems.* Solar Energy, 2001. **70**(4): p. 349-359.

224. Kim, S.-K., et al., *Modeling and simulation of a grid-connected PV generation system for electromagnetic transient analysis.* Solar Energy, 2009. **83**(5): p. 664-678.

225. Hernandez, J., G. Gordillo, and W. Vallejo, *Predicting the behavior of a grid-connected photovoltaic system from measurements of solar radiation and ambient temperature.* Applied Energy, 2013. **104**(0): p. 527-537.

226. Sukamongkol, Y., S. Chungpaibulpatana, and W. Ongsakul, *A simulation model for predicting the performance of a solar photovoltaic system with alternating current loads.* Renewable Energy, 2002. **27**(2): p. 237-258.

227. Hoff, T.E., R. Perez, and R.M. Margolis, *Maximizing the value of customer-sited PV systems using storage and controls.* Solar Energy, 2007. **81**(7): p. 940-945.

228. Nair, N.-K.C. and N. Garimella, *Battery energy storage systems: Assessment for small-scale renewable energy integration.* Energy and Buildings, 2010. **42**(11): p. 2124-2130.

229. Leadbetter, J. and L. Swan, *Battery storage system for residential electricity peak demand shaving.* Energy and Buildings, 2012. **55**(0): p. 685-692.

230. McKenna, E., et al., *Economic and environmental impact of lead-acid batteries in grid-connected domestic PV systems.* Applied Energy, 2013. **104**: p. 239-249.

231. Lijun, G., S. Liu, and R.A. Dougal, *Dynamic lithium-ion battery model for system simulation.* Components and Packaging Technologies, IEEE Transactions on, 2002. **25**(3): p. 495-505.

232. Darcovich, K., et al., *Higher-capacity lithium ion battery chemistries for improved residential energy storage with micro-cogeneration.* Applied Energy, 2013. **111**(0): p. 853-861.

233. West, J. and J.E. Braun, *Modeling partial charging and discharging of area-constrained ice storage tanks.* HVAC&R Research, 1999. **5**(3): p. 209-228.

234. Wilcox, S., *National Solar Radiation Database 1991-2005 Update: User's Manual*. 2007, National Renewable Energy Laboratory (NREL), Golden, CO.

235. Martin, N. and J. Ruiz, *Calculation of the PV modules angular losses under field conditions by means of an analytical model.* Solar Energy Materials and Solar Cells, 2001. **70**(1): p. 25-38.

236. Lu, L., *Investigation on Characteristics and Application of Hybrid Solar-Wind Power Generation Systems*. 2004, Ph.D. Dissertation, Hong Kong Polytechnic University: Hong Kong.

237. Chen, Z., *Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond filters, and beyond*. 2003, Adaptive Syst. Lab., McMaster Univ.,Adaptive Syst. Lab., McMaster Univ.,.

238. Bishop, G. and G. Welch, *An introduction to the kalman filter*. Proc of SIGGRAPH, Course, 2001. **8**: p. 27599-3175.

239. Juang, J.-N. and R.S. Pappa, *An eigensystem realization algorithm for modal parameter identification and model reduction.* Journal of guidance, control, and dynamics, 1985. **8**(5): p. 620-627.

240. Braun, J.E., *Load control using building thermal mass*. Journal of Solar Energy Engineering-Transactions of the Asme, 2003. **125**(3): p. 292-301.

Appendix A

Building Cluster Emulator Setup

1. EnergyPlus and BCVTB Connection

EnergyPlus and BCVTB are connected using Ptolemy through the external interface in EnergyPlus. Figure A-1and Figure A-2 show the external interface and the transferring variables.

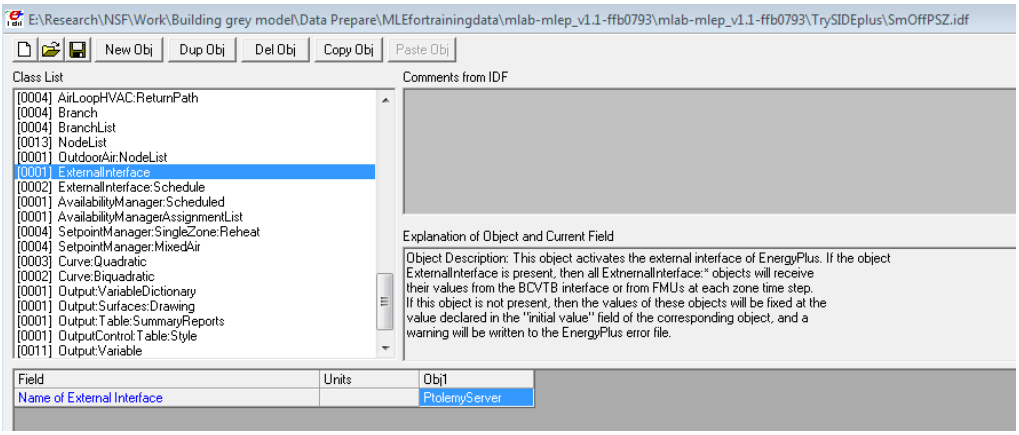


Figure A-1 External interface in EnergyPlus

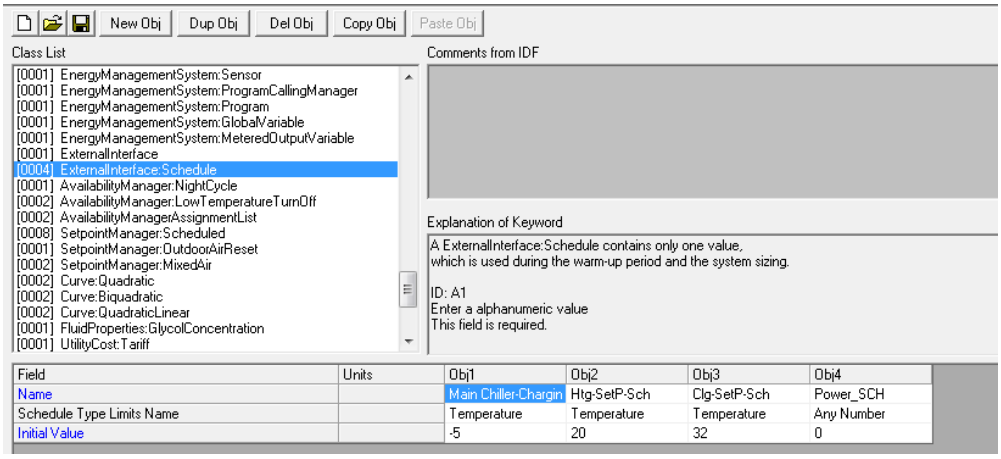


Figure A-2 Control variable transferring

In order to declare the data exchanging a configuration file is also needed:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<!DOCTYPE BCVTB-variables SYSTEM "variables.dtd">
<BCVTB-variables>
  <variable source="EnergyPlus">
   <EnergyPlus name="ENVIRONMENT" type="Outdoor Dry Bulb"/>
  </variable>
  <variable source="EnergyPlus">
    <EnergyPlus name="SPACE1-1" type="ZONE/SYS AIR TEMPERATURE"/>
  </variable>
  <variable source="Ptolemy">
    <EnergyPlus schedule="Htg-SetP-Sch"/>
  </variable>
  <variable source="Ptolemy">
    <EnergyPlus schedule="Clg-SetP-Sch"/>
  </variable>
  <variable source="Ptolemy">
    <EnergyPlus schedule="Power_SCH"/>
  </variable>
  <variable source="Ptolemy">
    <EnergyPlus schedule="Main Chiller-Charging Setpoint Temp
Schedule"/>
  </variable>

  <!-- The next two elements receive the schedule value as an output
from E+ -->
  <variable source="EnergyPlus">
    <EnergyPlus name="Htg-SetP-Sch" type="Schedule Value"/>
  </variable>
  <variable source="EnergyPlus">
    <EnergyPlus name="Clg-SetP-Sch" type="Schedule Value"/>
  </variable>
 <variable source="EnergyPlus">
    <EnergyPlus name="Power_SCH" type="Schedule Value"/>
  </variable>
</BCVTB-variables>
```

2. TRNSYS and BCVTB Connection

There is a MATLAB module in BCVTB for MATLAB, and TRNSYS and BCVTN are connected through MATLAB.

```
   dos('C:\TRNSYS17\Exe\TRNExe.exe
C:\TRNSYS17\MyProjects\Project1\Project1_test.dck /n')
   %% reading TRNSYS results
   fid=fopen('C:\TRNSYS17\MyProjects\Project1\Project1_test.out','rt');
   if fid==-1
```

```
      pout='Can not open sor file'
      pout='There is an error, please use control c to stop the
program'
      pause;
   end
   fgetl(fid);
   Trnsysdata   =   fscanf(fid,'%g%g%g%g%g%g%g%g%g%g%g%g%g%g%g%g%g%g%g',[3
inf]);
   % %Trnsysdata
   PV_power=Trnsysdata(2,2);

   %% update simulation time
   deltime=delTim/3600;
   path = 'C:\TRNSYS17\MyProjects\Project1\Project1_test.dck';
   FID = fopen(path,'rt');

   A = fscanf(FID,'%c');
   % Replace START time
   [str] = regexpi(A, '\w*START\s*=\s*\S*', 'match');
   str = str{1};
   pos = strfind(str,num2str('='));
   starttime_old = str2num(str(pos+1:end));
   starttime_new = starttime_old + deltime;
   str_new = [str(1:pos),' ',num2str(starttime_new)];
   A=strrep(A,str,str_new);

   % Replace STOP time
   [str] = regexpi(A, '\w*STOP\s*=\s*\S*', 'match');
   str = str{1};
   pos = strfind(str,num2str('='));
   stoptime_old = str2num(str(pos+1:end));
   stoptime_new = stoptime_old + deltime;
   str_new = [str(1:pos),' ',num2str(stoptime_new)];
   A=strrep(A,str,str_new);


   fclose(FID);

   FID = fopen(path,'wt');
   fprintf(FID,'%c',A); % Write out the whole file
   fclose(FID);
```

Appendix B

System Identification Codes

1. System excitation signal generation

```
clc; clear all; close all

L = 1;                      % time in days
n = 96;                     % number of points per period (15 mins)
t2 = linspace(0,L,n+1); t = t2(1:end-1)';

k=[0:(n/2-1) -n/2:-1]; % frequency components of FFT

Tmax = 32; % approx 90 F
Tmin = 10; % approx 50 F


dc = (Tmax - Tmin - 8) * rand + Tmin + 4;   % DC will be between Tmin +
4 and Tmax - 4

sig = dc * ones(96,1);

maxA = min([Tmax - dc, dc - Tmin]) / 4; % divide by 4 to insure some
high frequency content

for f = 1:96
    maxAarr(f,1) = maxA;
    omega = 2 * pi * f;
    A = maxA * rand;
    phi = 2 * pi * rand;
    sig = sig + A * sin(omega * t + phi);
    maxA = min([Tmax + 4 - max(sig), min(sig) - Tmin + 4]) / 4; %
divide by 2 to insure some high frequency content

subplot(3,1,1)
plot(24*t,sig)
xlabel(['Time (hours)'])
axis([0 24 10 35]); grid on
title(['DC is ' num2str(dc)])
subplot(3,1,2)
st = fft(sig);
plot(fftshift(k),abs(fftshift(st))/max(abs(fftshift(st))))
xlabel(['Wavenumber'])
axis([-48 48 0 1]); grid on
subplot(3,1,3)
st = fft(sig - dc * ones(96,1));
plot(fftshift(k),abs(fftshift(st))/max(abs(fftshift(st))))
xlabel('Wavenumber (no dc)')
axis([-48 48 0 1]); grid on
drawnow
```

```
end
figure
plot(24*t,sig)
xlabel(['Time (hours)'])
ylabel('Temperature Setpoint, C')
axis([0 24 10 35]); grid on
%title(['DC is ' num2str(dc)])
sigT=sig;
%% TEMPS generate a week's worth of inputs
clear all

clc;
L = 1;                      % time in days
n = 96;                     % number of points per period (15 mins)
t2 = linspace(0,L,n+1); t = t2(1:end-1)';
sigt=zeros(96,5);

k=[0:(n/2-1) -n/2:-1]; % frequency components of FFT

Tmax = 32; % approx 90 F
Tmin = 10; % approx 50 F

days = {'Monday','Tuesday','Wednesday','Thursday','Friday'};
for day = 1:5

    dc = (Tmax - Tmin - 8) * rand + Tmin + 4;   % DC will be between
Tmin + 4 and Tmax - 4

    sig = dc * ones(96,1);

    maxA = min([Tmax - dc, dc - Tmin]) / 4; % divide by 4 to insure
some high frequency content

    for f = 1:96
        maxAarr(f,1) = maxA;
        omega = 2 * pi * f;
        A = maxA * rand;
        phi = 2 * pi * rand;
        sig = sig + A * sin(omega * t + phi);
        maxA = min([Tmax + 4 - max(sig), min(sig) - Tmin + 4]) / 4; %
divide by 2 to insure some high frequency content
    end

    fprintf('    For: %s,\n', days{day})
    temps = round(10 * sig) / 10;
    %vect2EPsched(temps,15,'comma');
    sigt(:,day)=sig;
end
sigtt=sigt(:);

%% INFILTRATION and OCCUPANCY and LIGHTS and EQUIPMENT generate a
week's worth of inputs
```

```
clc;

L = 1;                      % time in days
n = 96;                     % number of points per period (15 mins)
t2 = linspace(0,L,n+1); t = t2(1:end-1)';

k=[0:(n/2-1) -n/2:-1]; % frequency components of FFT

smax = 1;
smin = 0;
sige=zeros(96,10);

days = {'Monday','Tuesday','Wednesday','Thursday','Friday','Starduay',
'Sunday','ScondMon','SecondTue','SecondWed'};
for day = 1:length(days)

    dc = (smax - smin) * 0.60 * rand + (smax - smin) * 0.20;   % DC
will be between Tmin + 4 and Tmax - 4

    sig = dc * ones(96,1);

    maxA = min([smax - dc, dc - smin]) / 4; % divide by 4 to insure
some high frequency content

    for f = 1:96
        maxAarr(f,1) = maxA;
        omega = 2 * pi * f;
        A = maxA * rand;
        phi = 2 * pi * rand;
        sig = sig + A * sin(omega * t + phi);
        maxA = min([smax + 0.2*(smax - smin) - max(sig), min(sig) -
smin + 0.2*(smax - smin)]) / 4; % divide by 2 to insure some high
frequency content
    end

    sig(sig > 1) = 1;
    sig(sig < 0) = 0;
    fprintf('    For: %s,\n', days{day})
    temps = round(100 * sig) / 100;
%   vect2EPsched(temps,15,'comma');
    sige(:,day)=sig;
end
%% LIGHTS and EQUIPMENT generate a week's worth of inputs

clc;

L = 1;                      % time in days
n = 96;                     % number of points per period (15 mins)
t2 = linspace(0,L,n+1); t = t2(1:end-1)';

k=[0:(n/2-1) -n/2:-1]; % frequency components of FFT

smax = 1;
smin = 0;
```

```
sigs=zeros(96,10);

days = {'Monday','Tuesday','Wednesday','Thursday','Friday','Starduay',
'Sunday','ScondMon','SecondTue','SecondWed'};
for day = 1:length(days)

    dc = (smax - smin) * 0.60 * rand + (smax - smin) * 0.20;   % DC
will be between Tmin + 4 and Tmax - 4

    sig = dc * ones(96,1);

    maxA = min([smax - dc, dc - smin]) / 4; % divide by 4 to insure
some high frequency content

    for f = 1:96
        maxAarr(f,1) = maxA;
        omega = 2 * pi * f;
        A = maxA * rand;
        phi = 2 * pi * rand;
        sig = sig + A * sin(omega * t + phi);
        maxA = min([smax + 0.2*(smax - smin) - max(sig), min(sig) -
smin + 0.2*(smax - smin)]) / 4; % divide by 2 to insure some high
frequency content
    end

    sig(sig > 1) = 1;
    sig(sig < 0) = 0;
    fprintf('    For: %s,\n', days{day})
    temps = round(100 * sig) / 100;
%   vect2EPsched(temps,15,'comma');
    sigs(:,day)=sig;
end
%%
subplot(3,1,1)
plot(24*t,sig)
xlabel(['Time (hours)'])
axis([0 24 0 1]); grid on
title(['DC is ' num2str(dc)])
subplot(3,1,2)
st = fft(sig);
plot(fftshift(k),abs(fftshift(st))/max(abs(fftshift(st))))
xlabel(['Wavenumber'])
axis([-48 48 0 1]); grid on
subplot(3,1,3)
st = fft(sig - dc * ones(96,1));
plot(fftshift(k),abs(fftshift(st))/max(abs(fftshift(st))))
xlabel('Wavenumber (no dc)')
axis([-48 48 0 1]); grid on
drawnow

figure
plot(24*t,sig)
xlabel(['Time (hours)'])
ylabel(['Equipment Schedule'])
```

```
axis([0 24 0 1]); grid on
%title(['DC is ' num2str(dc)])
sige=sig;

%% plotting
figure
[AX,H1,H2] = plotyy(24*t,sigt(:,1),24*t,sigs(:,1)*100,'plot');
set(get(AX(1),'Ylabel'),'String','Temperature')
set(get(AX(1),'Ylabel'),'String','Temperature Setpoint,C')
set(get(AX(2),'Ylabel'),'String','Equipment Schedule,%')
xlabel('Time, Hour');

%% Saving Eplus Inputs
%save('SIDEplusInput','sigT','sigs');

%%
figure
Tseche=zeros(1440,1);
Rseche=zeros(1440,1);
for i=1:48
    Tseche(30*(i-1)+1:30*i,1)=(sigt(2*i,1)+sigt(2*i-1,1))/2;
    Rseche(30*(i-1)+1:30*i,1)=(sigs(2*i,1)+sigs(2*i-1,1))/2;
end
L = 1;                      % time in days
n = 1440;                     % number of points per period (15 mins)
t3 = linspace(0,L,n+1); tt = t3(1:end-1)';

[AX,H1,H2] = plotyy(24*tt,Tseche,24*tt,Rseche*100,'plot');
set(get(AX(1),'Ylabel'),'String','Temperature')
set(get(AX(1),'Ylabel'),'String','Temperature Setpoint,C')
set(get(AX(2),'Ylabel'),'String','Equipment Schedule,%')
xlabel('Time, Hour');
axis(AX(1), [0 24 5 30])
axis(AX(2), [0 24 0 100])
%% Bandwith

fs = 1/1800;  % data collection at 1/1800 HZ
% pick a good time length. Because of the time shift we need
significantly more
% than 10s. Technically the sinc() has inifite length so we need to
extend
% it far enough to cover the vast part of the energy AND we need to
start
% at negative times, not at t = 0;
n = length (Tseche); % that's about 60s. Should be plenty
t = (-n/2:n/2-1)'/fs;   % time axis in seconds, symmetric around t = 0;

% frequency axis:
f0 = fs/n;  % frequency resoution in Hz
faxis = (-n/2:n/2-1)'*f0;   % frequency axis in Hz seconds, symmetric
around f = 0;
fx = fft(Tseche);
Pxx=circshift(abs(fx),n/2);
figure
```

```
plot(faxis,Pxx);
title('Spectrum of input signal');
xlabel('Frequency in Hz');


figure
plot(faxis(1:50),Pxx(1:50))
title('Power spectral density')
xlabel('Frequency (Hz)')
%% Crest factor
maxt=max(Tseche);
rmst = rms(Tseche);

CF_t=maxt/rmst;
fprintf('Crest Factor of the input signal: %f.\n', CF_t)

maxr=max(Rseche);
rmsr = rms(Rseche);
CF_r=maxr/rmsr;
fprintf('Crest Factor of the input signal: %f.\n', CF_r)
```

2. System characteristics test and model training

```
% Energy forecasting universial model
% Xiwang Li 10/09/2013

%% System identificaiton model traningin
function [M Y2 yhat mu_u mu_y]=EnergyForecastingTraining (u,y, dt)

r = size(u,1);  % number of inputs
m = size(y,1);  % number of outputs
N = length(u(1,:));
k = 1:N;
tt = dt * k;

mu_u = mean(u,2);
mu_y = mean(y,2);

std_u = std(u,[],2);
std_y = std(y,[],2);

u = bsxfun(@minus,u,mu_u);
y = bsxfun(@minus,y,mu_y);


Fs = 1 / dt; % in hour-1
recL = 12*Fs; % 6 hours
NFFT = 2048;
```

```
for i = 1:r
    for j = 1:i
        [P,~] = cpsd(u(i,:),u(j,:),recL,[],NFFT,Fs,'twosided');
        Suu(i,j,:) = P.';
        if j ~= i
            Suu(j,i,:) = P'; % conjugate transpose
        end
    end
end

%fprintf('Auto spectral power from time domain is %f\n',u*u'/length(u))
%fprintf('Auto spectral power from Suu is %f\n',(mean(Suu,3))*Fs)

Syu = zeros(m,r,NFFT);
for i = 1:m
    for j = 1:r
        [P,W] = cpsd(y(i,:),u(j,:),recL,[],NFFT,Fs,'twosided');
        Syu(i,j,:) = P.';
    end
end

G = zeros(m,r,NFFT);
for k = 1:NFFT
    G(:,:,k) = Syu(:,:,k)/Suu(:,:,k);
end

% Could lowpass filter the data, but does not seem to help
fil = 1./(1+1i*2*pi/100*W(1:NFFT/2));
fil = [fil; flipud(fil)];
%fil = ones(size(fil));

G2 = shiftdim(G,2);
%% Nonlinearity

Suu = zeros(r,r,NFFT);
Syy = zeros(m,r,NFFT);
for i = 1:r
    for j = 1:i
        [P,~] = cpsd(u(i,:),u(j,:),recL,[],NFFT,Fs,'twosided');
        Suu(i,j,:) = P.';
        if j ~= i
            Suu(j,i,:) = P'; % conjugate transpose
        end
    end
end

for i = 1:m
    for j = 1:i
        [Q,~] = cpsd(y(i,:),y(j,:),recL,[],NFFT,Fs,'twosided');
        Syy(i,j,:) = Q.';
        if j ~= i
            Syy(j,i,:) = Q'; % conjugate transpose
        end
    end
```

```
end

G = zeros(m,r,NFFT);
C = zeros(m,r,NFFT);
for k = 1:NFFT
    G(:,:,k) = Syu(:,:,k)/Suu(:,:,k);
    C(:,:,k) = (Syu(:,:,k)/Suu(:,:,k))*(Syu(:,:,k)/Syy(:,:,k));
 end

% Could lowpass filter the data, but does not seem to help
fil = 1./(1+1i*2*pi/100*W(1:NFFT/2));
fil = [fil; flipud(fil)];
%fil = ones(size(fil));

% Plot transfer functions
G2 = shiftdim(G,2);
C2 = shiftdim(C,2);
Cxy=mscohere(u,y,recL,[],NFFT,Fs,'twosided');

 figure
 plot(W(1:NFFT/2), Cxy(1:NFFT/2,1));
 %plot(Cxy,'DisplayName','Cxy','YDataSource','Cxy');
 axis ([0 6 0 1]);
 legend('SC_x_y'); xlabel('Frequency (hour^-^1)');
 ylabel('Magnitude')
 title('Coherence of the system');


%%
figure
subplot(2,3,1)
 plot(W(1:NFFT/2)/6, abs(C2(1:NFFT/2,:,1)),'k','LineWidth',3);
title('Tout','fontweight','normal', 'FontSize', 16); axis ([0 1 0 1]);
ylabel('Nonlinearity Index','fontweight','bold','FontSize', 20);
subplot(2,3,2)
 plot(W(1:NFFT/2)/6, abs(C2(1:NFFT/2,:,2)),'k','LineWidth',3);
title('Tsep','fontweight','normal', 'FontSize', 16); axis ([0 1 0 1]);
subplot(2,3,3)
 plot(W(1:NFFT/2)/6, abs(C2(1:NFFT/2,:,11)),'k','LineWidth',3);
title('Qdir', 'fontweight','normal', 'FontSize', 16);  axis ([0 1 0 1]);
subplot(2,3,4)
 plot(W(1:NFFT/2)/6, abs(C2(1:NFFT/2,:,12)),'k','LineWidth',3);
title('Qdif','fontweight','normal', 'FontSize', 16); xlabel('Normalized
Frequency (h^-^1)','fontweight','bold','FontSize', 16);axis ([0 1 0 1]);
ylabel('Nonlinearity Index','fontweight','bold','FontSize', 20);
subplot(2,3,5)
 plot(W(1:NFFT/2)/6, abs(C2(1:NFFT/2,:,8)),'k','LineWidth',3);
title('Rin','fontweight','normal', 'FontSize', 16); xlabel('Normalized
Frequency (h^-^1)','fontweight','bold','FontSize', 16);axis ([0 1 0 1]);
 subplot(2,3,6)
 plot(W(1:NFFT/2)/6, abs(C2(1:NFFT/2,:,end)),'k','LineWidth',3);
title('Voa','fontweight','normal', 'FontSize', 16); xlabel('Normalized
Frequency (h^-^1)','fontweight','bold','FontSize', 16);axis ([0 1 0 1]);
```

```
%% Calculate squared Coherence

CxyTout=mscohere(u(1,:),y,recL,[],NFFT,Fs,'twosided'); % Tout
CxyTcore=mscohere(u(2,:),y,recL,[],NFFT,Fs,'twosided');
CxyTsout=mscohere(u(3,:),y,recL,[],NFFT,Fs,'twosided');
CxyTwest=mscohere(u(4,:),y,recL,[],NFFT,Fs,'twosided');
CxyTnorth=mscohere(u(5,:),y,recL,[],NFFT,Fs,'twosided');
CxyTeast=mscohere(u(6,:),y,recL,[],NFFT,Fs,'twosided');
CxyTdirect=mscohere(u(11,:),y,recL,[],NFFT,Fs,'twosided');
CxyTdiff=mscohere(u(12,:),y,recL,[],NFFT,Fs,'twosided');
CxyTequip=mscohere(u(8,:),y,recL,[],NFFT,Fs,'twosided');
CxyTventi=mscohere(u(end,:),y,recL,[],NFFT,Fs,'twosided');


figure
subplot(2,5,1)
hist(CxyTout);  title('Tout','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]);ylabel('count
#','fontweight','normal','FontSize', 20);
subplot(2,5,2)
hist(CxyTcore);  title('Tcore,stp','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]);
subplot(2,5,3)
hist(CxyTsout);  title('Tsouth,stp','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]);
subplot(2,5,4)
hist(CxyTwest);  title('Twest,stp','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]);
subplot(2,5,5)
hist(CxyTnorth);  title('Tnorth,stp','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]);
subplot(2,5,6)
hist(CxyTeast);  title('Teast,stp','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]);xlabel('Normalnized Frequency (h^-
^1)','fontweight','normal','FontSize', 16);ylabel('count
#','fontweight','normal','FontSize', 20);
subplot(2,5,7)
hist(CxyTdirect);  title('Qdir','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]);xlabel('Normalnized Frequency (h^-
^1)','fontweight','normal','FontSize', 16);
subplot(2,5,8)
hist(CxyTdiff);  title('Qdif','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]); xlabel('Normalnized Frequency (h^-
^1)','fontweight','normal','FontSize', 16);
subplot(2,5,9)
hist(CxyTequip);  title('Rin','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]); xlabel('Normalnized Frequency (h^-
^1)','fontweight','normal','FontSize', 16);
subplot(2,5,10)
hist(CxyTventi);  title('Voa','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]); xlabel('Normalnized Frequency (h^-
^1)','fontweight','normal','FontSize', 16);
```

```
figure
subplot(2,5,1)
 plot(W(1:NFFT/2)/6, CxyTout(1:NFFT/2,1),'k','LineWidth',3);
title('Tout','fontweight','normal', 'FontSize', 16); axis ([0 1 0 1]);
ylabel('Nonlinearity Index','fontweight','bold','FontSize', 20);
subplot(2,5,2)
 plot(W(1:NFFT/2)/6, CxyTcore(1:NFFT/2,1),'k','LineWidth',3);
title('Tcore,stp','fontweight','normal', 'FontSize', 16); axis ([0 1 0
1]);
subplot(2,5,3)
 plot(W(1:NFFT/2)/6, CxyTsout(1:NFFT/2,1),'k','LineWidth',3);
title('Tsouth,stp','fontweight','normal', 'FontSize', 16); axis ([0 1 0
1]);
subplot(2,5,4)
 plot(W(1:NFFT/2)/6, CxyTwest(1:NFFT/2,1),'k','LineWidth',3);
title('Twest','fontweight','normal', 'FontSize', 16); axis ([0 1 0 1]);
subplot(2,5,5)
 plot(W(1:NFFT/2)/6, CxyTnorth(1:NFFT/2,1),'k','LineWidth',3);
title('Tnorth,stp','fontweight','normal', 'FontSize', 16); axis ([0 1 0
1]);
subplot(2,5,6)
 plot(W(1:NFFT/2)/6, CxyTeast(1:NFFT/2,1),'k','LineWidth',3);
title('Teast,stp','fontweight','normal', 'FontSize', 16);
xlabel('Normalnized Frequency (h^-^1)','fontweight','bold','FontSize',
16);axis ([0 1 0 1]);  ylabel('Nonlinearity
Index','fontweight','bold','FontSize', 20);
subplot(2,5,7)
 plot(W(1:NFFT/2)/6, CxyTdirect(1:NFFT/2,1),'k','LineWidth',3);
title('Qdir','fontweight','normal', 'FontSize', 16);
xlabel('Normalnized Frequency (h^-^1)','fontweight','bold','FontSize',
16);axis ([0 1 0 1]);
subplot(2,5,8)
 plot(W(1:NFFT/2)/6, CxyTdiff(1:NFFT/2,1),'k','LineWidth',3);
title('Qdif','fontweight','normal', 'FontSize', 16);
xlabel('Normalnized Frequency (h^-^1)','fontweight','bold','FontSize',
16);axis ([0 1 0 1]);
subplot(2,5,9)
 plot(W(1:NFFT/2)/6, CxyTequip(1:NFFT/2,1),'k','LineWidth',3);
title('Rin','fontweight','normal', 'FontSize', 16); xlabel('Normalnized
Frequency (h^-^1)','fontweight','bold','FontSize', 16);axis ([0 1 0 1]);
subplot(2,5,10)
 plot(W(1:NFFT/2)/6, CxyTventi(1:NFFT/2,1),'k','LineWidth',3);
title('Voa','fontweight','normal', 'FontSize', 16); xlabel('Normalnized
Frequency (h^-^1)','fontweight','bold','FontSize', 16);axis ([0 1 0 1]);

 figure
subplot(2,3,1)
 plot(W(1:NFFT/2), CxyTout(1:NFFT/2,1),'k','LineWidth',6); title('Out
Temp','fontweight','bold', 'FontSize', 24); axis ([0 6 0 1]);
 subplot(2,3,2)
 plot(W(1:NFFT/2), abs(C2(1:NFFT/2,:,7)),'k','LineWidth',6);
title('Plenum Temp','fontweight','bold', 'FontSize', 24); axis ([0 6 0
1]);
```

```
subplot(2,3,3)
 plot(W(1:NFFT/2), abs(C2(1:NFFT/2,:,2)),'k','LineWidth',6);
title('Zone Tsetp','fontweight','bold', 'FontSize', 24); axis ([0 6 0
1]);
subplot(2,3,4)
 plot(W(1:NFFT/2), abs(C2(1:NFFT/2,:,11)),'k','LineWidth',6);
title('Direct Solar', 'fontweight','bold','FontSize', 24);
xlabel('Normalnized Frequency (h^-^1)','fontweight','bold','FontSize',
30);axis ([0 6 0 1]);
subplot(2,3,5)
 plot(W(1:NFFT/2), abs(C2(1:NFFT/2,:,12)),'k','LineWidth',6);
title('Diffuse Solar','fontweight','bold', 'FontSize', 24);
xlabel('Normalnized Frequency (h^-^1)','fontweight','bold','FontSize',
30);axis ([0 6 0 1]);
subplot(2,3,6)
 plot(W(1:NFFT/2), CxyTequip(1:NFFT/2,1),'k','LineWidth',6); title('Lg
& Eq Sch','fontweight','bold','FontSize', 24); xlabel('Normalnized
Frequency (h^-^1)','fontweight','bold','FontSize', 30);axis ([0 6 0 1]);
%%
% Calculate Markov parameters
Y = zeros(size(G));
for i=1:m
    for j = 1:r
        Y(i,j,:) = ifft(G2(:,:,j));
    end
end
% Plot Markov parameters
last = min(recL,NFFT);
Y2 = shiftdim(Y,2);
%% plots of markov parameter (coefficient of each input in transfer
function)
figure
subplot(2,5,1)
plot(tt(1:last),Y2(1:last,:,1),'x-'); title('Core Zone Temp');
xlabel('Normalnized Frequency (h^-^1)');
subplot(2,5,2)
plot(tt(1:last),Y2(1:last,:,2),'x-');title('South Zone Temp');
xlabel('Normalnized Frequency (h^-^1)');
subplot(2,5,3)
plot(tt(1:last),Y2(1:last,:,3),'x-'); title('West Zone Temp');
xlabel('Normalnized Frequency (h^-^1)');
subplot(2,5,4)
plot(tt(1:last),Y2(1:last,:,4),'x-'); title('North Zone Temp');
xlabel('Normalnized Frequency (h^-^1)');
subplot(2,5,5)
plot(tt(1:last),Y2(1:last,:,5),'x-'); title('East Zone Temp');
xlabel('Normalnized Frequency (h^-^1)');
subplot(2,5,6)
plot(tt(1:last),Y2(1:last,:,1),'x-'); title('Outside Temp');
xlabel('Normalnized Frequency (h^-^1)');
subplot(2,5,7)
plot(tt(1:last),Y2(1:last,:,11),'x-'); title('Direct Solar Radi');
xlabel('Normalnized Frequency (h^-^1)');
subplot(2,5,8)
```

```
plot(tt(1:last),Y2(1:last,:,12),'x-'); title('Diffuse Solar Radi');
xlabel('Normalnized Frequency (h^-^1)');
subplot(2,5,9)
plot(tt(1:last),Y2(1:last,:,8),'x-');title('Equip Schedule');
xlabel('Normalnized Frequency (h^-^1)');
subplot(2,5,10)
plot(tt(1:last),Y2(1:last,:,end),'x-'); title('Ventilation');
xlabel('Normalnized Frequency (h^-^1)');

%only use the Markov parameters from first half (record length)
%%
Y2 = Y2(1:last,:,:);
% Calculate estimates
yhat = zeros(size(y));
for i=1:m
    for j = 1:r
        resp = conv(Y2(:,i,j),u(j,:));
        yhat(i,:) = yhat(i,:) + resp(1:length(yhat));
    end
end

err = yhat - y;
stderr = std(err);
VAF = 100*(1-(stderr/std_y)^2);
y = y + mu_y;
yhat = yhat + mu_y;
save('modelFinal','mu_u','mu_y','std_u','std_y','Y2');
M = shiftdim(Y2,2);
save('simpleModelParams.mat','mu_u','mu_y','M');
fprintf('\nThe variance accounted for by the model is %4.2f%%\n\n',VAF)

end
```

3. System identification model forecasting

```
% Energy forecasting universal model Forecasting
% Xiwang Li 10/09/2013

function [yhat]=EnergyForcastingFinal (u,y,Y2,mu_u, mu_y)

dt=1/12;

Fs = 1 / dt; % in hour-1
recL = 12*Fs; % 6 hours
NFFT = 2048;


CxyTout=mscohere(u(1,:),y,recL,[],NFFT,Fs,'twosided'); % Tout
CxyTcore=mscohere(u(2,:),y,recL,[],NFFT,Fs,'twosided');
CxyTsout=mscohere(u(3,:),y,recL,[],NFFT,Fs,'twosided');
CxyTwest=mscohere(u(4,:),y,recL,[],NFFT,Fs,'twosided');
CxyTnorth=mscohere(u(5,:),y,recL,[],NFFT,Fs,'twosided');
```

```
CxyTeast=mscohere(u(6,:),y,recL,[],NFFT,Fs,'twosided');
CxyTdirect=mscohere(u(11,:),y,recL,[],NFFT,Fs,'twosided');
CxyTdiff=mscohere(u(12,:),y,recL,[],NFFT,Fs,'twosided');
CxyTequip=mscohere(u(8,:),y,recL,[],NFFT,Fs,'twosided');
CxyTventi=mscohere(u(end,:),y,recL,[],NFFT,Fs,'twosided');

figure
subplot(2,5,1)
hist(CxyTout);  title('Tout','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]);ylabel('count
#','fontweight','normal','FontSize', 20);
subplot(2,5,2)
hist(CxyTcore);  title('Tcore,stp','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]);
subplot(2,5,3)
hist(CxyTsout);  title('Tsouth,stp','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]);
subplot(2,5,4)
hist(CxyTwest);  title('Twest,stp','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]);
subplot(2,5,5)
hist(CxyTnorth);  title('Tnorth,stp','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]);
subplot(2,5,6)
hist(CxyTeast);  title('Teast,stp','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]);xlabel('Normalnized Frequency (h^-
^1)','fontweight','normal','FontSize', 16);ylabel('count
#','fontweight','normal','FontSize', 20);
subplot(2,5,7)
hist(CxyTdirect);  title('Qdir','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]);xlabel('Normalnized Frequency (h^-
^1)','fontweight','normal','FontSize', 16);
subplot(2,5,8)
hist(CxyTdiff);  title('Qdif','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]); xlabel('Normalnized Frequency (h^-
^1)','fontweight','normal','FontSize', 16);
subplot(2,5,9)
hist(CxyTequip);  title('Rin','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]); xlabel('Normalnized Frequency (h^-
^1)','fontweight','normal','FontSize', 16);
subplot(2,5,10)
hist(CxyTventi);  title('Voa','fontweight','normal', 'FontSize',
16);axis ([0 0.8 0 2000]); xlabel('Normalnized Frequency (h^-
^1)','fontweight','normal','FontSize', 16);

r = size(u,1);  % number of inputs
m = size(y,1);  % number of outputs


std_u = std(u,[],2);
std_y = std(y,[],2);

u = bsxfun(@minus,u,mu_u);
```

```
y = bsxfun(@minus,y,mu_y);


yhat = zeros(size(y));

 for i=1:m
    for j = 1:r
        resp = conv(Y2(:,i,j),u(j,:));
        %response(j,:) = resp(1:length(yhat));
        yhat(i,:) = yhat(i,:) + resp(1:length(yhat));
    end
 end

yhat = yhat + mu_y;
y=y + mu_y;


err = yhat - y;
stderr = std(err);
VAF = 100*(1-(stderr/std_y)^2);

fprintf('\nThe variance accounted for by the model training
is %4.2f%%\n\n',VAF)

 for i=1:length(yhat)
    if (yhat(i)<=0)
        yhat(i)=0;
    end
 end
end
```

4. State Space model development and Kalman Filter implementation

```
% Save and reading all Markov parameters for state space model
realization
load('E:\Research\NSF\Work\Building grey model\Testing ID\Code
achieve\Markov\Core_cooling\modelFinal.mat')
%ERA
r = 10;
m = 1;
% Note: when forming the Hankel matrix for SISO system if tthe length
of
% included Markov parameters is b, the biggest full square size matrix
is
% ceil(b/2). Code below assumes b is an even number
M = shiftdim(Y2,2);
btemp = size(M,2);
MM = M(:)';
```

```
Htemp = zeros(btemp*m,btemp*r);
for i = 1:btemp
    Htemp(i,1:btemp*r-(i-1)*r) = MM((i-1)*r+1:end);
end
b =35; % determine alpa and beta
H0 = Htemp(1:b*m,1:b*r);
H1 = Htemp(1:b*m,r+1:(b+1)*r);

[R,Sigma,S] = svd(H0); % H0 = R*Sigma*S' , using Juang notation

fprintf('First ten singular values of H0');
diag(Sigma);

fprintf('Use order 4\n');
n = 5;

Rn = R(:,1:n);
Sn = S(:,1:n);
Sigma_n = Sigma(1:n,1:n);

fprintf('For n = %d, the 2-norm of (Rn*Sigma_n*Sn'' - H0) is %f\n', n,
norm(Rn*Sigma_n*Sn' - H0));

Im = eye(m);
Om = zeros(m);
Em = [Im; repmat(Om,b*m-m,1)];
Ir = eye(r);
Or = zeros(r);
%Er = [Ir; repmat(Or,b/2-r,1)];  % it seems Juang should not have
transpose
Er = [Ir; zeros(b*r-r,r)];

fprintf('The estimated system is\n')
Ahat = (Sigma_n^-0.5)*Rn'*H1*Sn*(Sigma_n^-0.5);
Bhat = (Sigma_n^0.5)*Sn'*Er;
Chat = Em'*Rn*(Sigma_n^0.5);

save('stateSpace.mat','Ahat','Bhat','Chat')


 %% Form the static data from the baseline

deltEP = 5;     % simulation (energyplus) timestep in minutes
deltReg = 15;   % regression timestep in minutes
dt = deltReg/60/24; % dt is in days
cr = deltReg / deltEP;     % compression ratio
bldgInfo = defineBldg('SmallOffice');
zones = bldgInfo.zones;
EPout_realinput = loadEPdata_noexcite(bldgInfo,deltEP,'OneYear');

RegData_realinput = compressEPtoReg_energy(EPout_realinput,cr,zones);
```

```
clear EPout deltEP N_EP

load('E:\Research\NSF\Work\Building grey model\Testing ID\Code
achieve\Markov\Core_cooling\modelFinal.mat')

y=RegData_realinput.CORE_ZN(:,11)'./300;

%y=[y(:,289:432) y(:,289:432) y(:,289:432)];  % Heating Coil Electric
Energy,J
u1 = RegData_realinput.CORE_ZN(:,1)';         % Tzone at end of
timestep
u2 = RegData_realinput.CORE_ZN(:,2)';         % Qgsic average during
timestep
u3 = RegData_realinput.CORE_ZN(:,3)';         % Qgsir average during
timestep
u4 = RegData_realinput.ATTIC(:,1)';           % Tzone average during
timestep
u5 = RegData_realinput.PERIMETER_ZN_1(:,1)';   % Tzone average during
timestep
u6 = RegData_realinput.PERIMETER_ZN_2(:,1)';   % Tzone average during
timestep
u7 = RegData_realinput.PERIMETER_ZN_3(:,1)';   % Tzone average during
timestep
u8 = RegData_realinput.PERIMETER_ZN_4(:,1)';   % Tzone average during
timestep
u9 = RegData_realinput.outdoor(:,1)';          % Toa average during
timestep
u10 = RegData_realinput.CORE_ZN (:,6)';  % Sensible Heating Rate
u11 = RegData_realinput.CORE_ZN(:,12)'./300;  % Fan heat injection,J

%u1=18*ones(size(u2));
u = [u1; u2; u3;u4; u5; u6; u7;u8;u9;u11];
save('staticData.mat','y','u1','u');


%% Run the model

load('stateSpace.mat');

%load('staticData.mat');
n = size(Ahat,2);
u = bsxfun(@minus,u,mu_u);
y = bsxfun(@minus,y,mu_y);
yhat = zeros(size(y));
%u = [u1 ; u];
K = size(y,2);
x = zeros(n,K+1);
x(:,1) = zeros(n,1);

for k = 1:K
    x(:,k+1) = Ahat*x(:,k) + Bhat*u(:,k);
```

```
    yhat(k+1) = Chat*x(:,k+1);
    yhat(k+1) = Chat*x(:,k+1) + mu_y;
end
y = y + mu_y;
yhat = yhat + mu_y;

%% system off
yhat(1, 1:20)=0;
yhat(1, 86:116)=0;
yhat(1, 181:212)=0;
yhat(1, 277:288)=0;
%%
figure
plot(dt*(1:length(y)),yhat(1:288),'bo',dt*(1:length(y)),y,'rx'); grid
on
legend('Ees','Eep')
axis ([0 3 -1000 3000]);
xlabel('Time (days)'); ylabel('Core Zone Cooling Energy (W)')
title('Core Zone Cooling Energy, W');

figure
plot(dt*(1:length(y)),yhat(1:288),'r',dt*(1:length(y)),y,'k'); grid on
legend('Ees','Eep')
axis ([0 3 -1000 3000]);
xlabel('Time (days)'); ylabel('Core Zone Cooling Energy (W)')
title('Core Zone Cooling Energy, W');

[r2 rmse] = rsquare(y,yhat(1:288));
fprintf('\nThe R2 for the model forecasting is %4.2f%%\n\n',r2*100)

%% Kalman filter
% simulate the actual system with noise and create noisy measurement
vector
nk = 288;   % set this for the number of times to loop
state_p_var=100;  % process noise from ERA model
meas_var=10000;     % measurement noise
s1 = rng(6777);  % set up so we get same random numbers each time
w = sqrt(state_p_var)* randn(1,nk);
s2 = rng(1234); % different seed for measurements
r=normrnd(0,sqrt(meas_var),nk,1);
% noisy for measurement
F=Ahat;
G=Bhat;
H=Chat;

%  initialization of Kalman filter
y_est = zeros(1,nk); % clear out vector
y_meas=y+r;
xkp = zeros(n,1);
Pkp = 100*eye(n);  % initial covariance matrix

Q=eye(n)*state_p_var;
```

```
Q(n,n)=10e5;
R=meas_var;

xc=xkp;
K_acc = [];
innov = zeros(n,n);

for k = 1:nk

    Pkm = F*Pkp*F' + Q;
    Kk = Pkm*(H'/(H*Pkm*H' + R));
    xkm=F*xkp+G*u(:,k);
   if k<2
        yk=y_meas(:,k);
   else
       yk = y_meas(:,k);
   end
    innv = (yk -H*xkm);
    innov(:,k) = innv;
    xkp = xkm + Kk*innv;
    y_est(k)=H*xkp;
    xc = [xc,xkp];
    Pkp = (eye(n) - Kk*H)*Pkm*(eye(n) - Kk*H)' + Kk*R*Kk';
    error(k)=(y(k)- y_est(k));
end

figure
plot(dt*(1:length(y)),y_est,'bo',dt*(1:length(y)),y,'rx'); grid on
legend('Ees','Eep')
axis ([0 3 -1000 3000]);
xlabel('Time (days)'); ylabel('Core Zone Cooling Energy (W)')
title('Core Zone Cooling Energy, W');

figure
plot(dt*(1:length(y)),y_est,'r',dt*(1:length(y)),y,'k',dt*(1:length(y))
,y_meas+r,'g'); grid on
legend('Ees','Eep', 'Emea')
axis ([0 3 -1000 3000]);
xlabel('Time (days)'); ylabel('Core Zone Cooling Energy (W)')
title('Core Zone Cooling Energy, W');


[r2 rmse] = rsquare(y,y_est);
fprintf('\nThe R2 for the KF model forecasting is %4.2f%%\n\n',r2)


err = y - y_est;
stderr = std(err);
VAF = 100*(1-(stderr/std_y)^2);
fprintf('\nThe R2 for the KF model forecasting is %4.2f%%\n\n',VAF)
```

Appendix C

Experiment Design

This experiment will be conducted at Iowa Energy Center, Energy Resource Station for two weeks in August (14 days). The experiment will use system A and system B for VAV system. The two systems (A and B) will run all the tests to represent two different floors.

This experiment has two tasks:

1. System characteristics test
2. System excitation and identification

The system characteristics test will test the system nonlinearity and system response time. The information form this test will be used to generate the system excitation signals.

Objective

The objective of this experiment is to test the system identification methodology for building energy forecasting. This methodology has been tested and validated on simulation model. This experiment will apply and verify the performance of this methodology on real building.

Detailed Plan

1. Time management

	M	T	W	R	F	S	Sun
Date	Day 1	Day 2	Day 3	Day 4	Day 5	Day 6	Day 7
Test	Nonlinearity test	Nonlinearity test	Response time test	Normal	System excitation		
Date	Day 8	Day 9	Day 10	Day 11	Day 12	Day 13	Day 14
Test	Normal		System excitation		Normal	System excitation	Normal

2. Nonlinearity test

Control variables: During the testing period, the room temperature setpoint and equipment schedules will be changed according to a predetermined plan (every 30 minutes).

The temperature setpoints during the nonlinearity test days (Day 1 and Day 2) are listed in Table C1. The equipment schedules are summarized in Table C1. For the equipment schedule, there are 6 stages for lighting and 2 stages for the base board heat. Considering the base board has much higher power than the lighting, 6 stages have been chosen for the excitation.

Stage zero: all off

Stage one: lighting stage 1 on only

Stage two: lighting stage 2 on only

Stage three: lighting stage 1 and 2 on

Stage four: Base board 1 on

Stage Five: Base board 1 and 2 on

Stage Six: Everything on

Table C1. Nonlinearity Test Signal: Temperature Setpoints for system A

Day 1			Day 2	
Time	Cooling setpoint, F	Heating Setpoint, F	Cooling setpoint, F	Heating Setpoint, F
0:00	77.9	72.5	75.4	70.0
0:30	69.6	64.2	72.9	67.5
1:00	64.2	58.8	70.2	64.8
1:30	71.4	66.0	65.7	60.3
2:00	66.6	61.2	68.4	63.0
2:30	70.0	64.6	70.7	65.3
3:00	73.8	68.4	74.1	68.7
3:30	70.9	65.5	66.0	60.6
4:00	70.9	65.5	74.5	69.1
4:30	69.1	63.7	79.7	74.3
5:00	59.7	54.3	68.0	62.6
5:30	71.2	65.8	71.4	66.0
6:00	73.4	68.0	67.5	62.1
6:30	65.1	59.7	67.1	61.7
7:00	64.4	59.0	71.2	65.8
7:30	72.0	66.6	66.9	61.5
8:00	64.4	59.0	70.3	64.9
8:30	63.7	58.3	77.0	71.6
9:00	72.7	67.3	72.7	67.3
9:30	74.3	68.9	73.6	68.2
10:00	65.1	59.7	78.3	72.9
10:30	67.6	62.2	74.1	68.7
11:00	70.7	65.3	71.2	65.8
11:30	71.8	66.4	76.8	71.4
12:00	71.4	66.0	66.6	61.2
12:30	68.2	62.8	67.3	61.9
13:00	71.1	65.7	77.0	71.6
13:30	69.6	64.2	71.4	66.0
14:00	79.0	73.6	69.6	64.2
14:30	71.4	66.0	71.4	66.0
15:00	82.2	76.8	68.9	63.5
15:30	72.1	66.7	68.7	63.3
16:00	66.7	61.3	68.4	63.0

16:30	73.0	67.6	67.8	62.4
17:00	63.3	57.9	66.4	61.0
17:30	75.4	70.0	69.6	64.2
18:00	72.9	67.5	67.8	62.4
18:30	70.2	64.8	75.0	69.6
19:00	65.7	60.3	70.7	65.3
19:30	74.5	69.1	71.6	66.2
20:00	79.7	74.3	71.1	65.7
20:30	68.0	62.6	74.3	68.9
21:00	71.4	66.0	73.8	68.4
21:30	67.5	62.1	77.7	72.3
22:00	63.7	58.3	75.4	70.0
22:30	70.2	64.8	73.0	67.6
23:00	68.9	63.5	71.2	65.8
23:30	64.6	59.2	67.8	62.4
0:00	59.4	54.0	69.4	64.0

Table C2. Nonlinearity Test Signal: Equipment Schedule for system A

	Day 1			Day 2		
Time	Stage	Lighting	Base board	Stage	Lighting	Base board
0:00	0	off	off	0	off	off
0:30	5	stage 1 and 2	Stage 1	1	Stage 1	off
1:00	0	off	off	4	off	Stage 1
1:30	4	off	Stage 1	0	off	off
2:00	5	stage 1 and 2	Stage 1	0	off	off
2:30	2	stage 2	off	2	stage 2	off
3:00	6	off	Stage 2	0	off	off
3:30	3	stage 1 and 2	off	1	Stage 1	off
4:00	0	off	off	0	off	off
4:30	5	stage 1 and 2	Stage 1	1	Stage 1	off
5:00	5	stage 1 and 2	Stage 1	4	off	Stage 1
5:30	2	stage 2	off	2	stage 2	off
6:00	4	off	Stage 1	2	stage 2	off
6:30	3	stage 1 and 2	off	5	stage 1 and 2	Stage 1
7:00	4	off	Stage 1	1	Stage 1	off
7:30	3	stage 1 and 2	off	5	stage 1 and 2	Stage 1
8:00	3	stage 1 and 2	off	6	off	Stage 2

8:30	2	stage 2	off	4	off	Stage 1
9:00	5	stage 1 and 2	Stage 1	2	stage 2	off
9:30	5	stage 1 and 2	Stage 1	3	stage 1 and 2	off
10:00	6	off	Stage 2	2	stage 2	off
10:30	2	stage 2	off	0	off	off
11:00	5	stage 1 and 2	Stage 1	2	stage 2	off
11:30	2	stage 2	off	0	off	off
12:00	5	stage 1 and 2	Stage 1	2	stage 2	off
12:30	2	stage 2	off	5	stage 1 and 2	Stage 1
13:00	2	stage 2	off	2	stage 2	off
13:30	4	off	Stage 1	4	off	Stage 1
14:00	1	Stage 1	off	2	stage 2	off
14:30	0	off	off	6	off	Stage 2
15:00	6	off	Stage 2	5	stage 1 and 2	Stage 1
15:30	5	stage 1 and 2	Stage 1	5	stage 1 and 2	Stage 1
16:00	5	stage 1 and 2	Stage 1	2	stage 2	off
16:30	3	stage 1 and 2	off	6	off	Stage 2
17:00	4	off	Stage 1	2	stage 2	off
17:30	0	off	off	6	off	Stage 2
18:00	5	stage 1 and 2	Stage 1	3	stage 1 and 2	off
18:30	1	Stage 1	off	3	stage 1 and 2	off
19:00	6	off	Stage 2	6	off	Stage 2
19:30	2	stage 2	off	5	stage 1 and 2	Stage 1
20:00	4	off	Stage 1	5	stage 1 and 2	Stage 1
20:30	1	Stage 1	off	6	off	Stage 2
21:00	2	stage 2	off	0	off	off
21:30	5	stage 1 and 2	Stage 1	2	stage 2	off
22:00	6	off	Stage 2	5	stage 1 and 2	Stage 1
22:30	5	stage 1 and 2	Stage 1	2	stage 2	off
23:00	0	off	off	6	off	Stage 2
23:30	0	off	off	5	stage 1 and 2	Stage 1
0:00	2	stage 2	off	5	stage 1 and 2	Stage 1

Table C3. Nonlinearity Test Signal: Temperature Setpoints for system B

Day 1			Day 2	
Time	Cooling setpoint, F	Heating Setpoint, F	Cooling setpoint, F	Heating Setpoint, F
0:00	65.9	60.9	67.6	62.6
0:30	76.4	71.4	65.5	60.5
1:00	65.4	60.4	63.0	58.0
1:30	69.6	64.6	68.3	63.3
2:00	68.6	63.6	72.6	67.6
2:30	65.7	60.7	67.2	62.2
3:00	65.1	60.1	75.5	70.5
3:30	72.1	67.1	66.6	61.6
4:00	64.8	59.8	77.5	72.5
4:30	65.0	60.0	67.3	62.3
5:00	73.9	68.9	70.2	65.2
5:30	69.8	64.8	75.0	70.0
6:00	71.2	66.2	74.2	69.2
6:30	66.6	61.6	67.1	62.1
7:00	69.1	64.1	72.5	67.5
7:30	70.9	65.9	65.8	60.8
8:00	70.4	65.4	63.7	58.7
8:30	76.2	71.2	72.2	67.2
9:00	67.6	62.6	63.6	58.6
9:30	69.3	64.3	62.1	57.1
10:00	72.8	67.8	69.6	64.6
10:30	73.2	68.2	74.8	69.8
11:00	62.7	57.7	74.2	69.2
11:30	68.4	63.4	64.4	59.4
12:00	67.4	62.4	69.9	64.9
12:30	69.7	64.7	72.7	67.7
13:00	71.7	66.7	73.3	68.3
13:30	71.4	66.4	75.4	70.4
14:00	73.9	68.9	75.5	70.5
14:30	71.4	66.4	72.9	67.9
15:00	70.9	65.9	64.7	59.7
15:30	73.2	68.2	69.9	64.9
16:00	71.3	66.3	64.8	59.8
16:30	65.0	60.0	74.2	69.2

17:00	70.2	65.2	76.1	71.1
17:30	63.8	58.8	62.6	57.6
18:00	73.8	68.8	72.8	67.8
18:30	62.2	57.2	68.6	63.6
19:00	67.8	62.8	68.1	63.1
19:30	70.1	65.1	66.6	61.6
20:00	69.7	64.7	74.4	69.4
20:30	76.6	71.6	71.2	66.2
21:00	70.0	65.0	69.4	64.4
21:30	67.6	62.6	71.3	66.3
22:00	71.9	66.9	73.5	68.5
22:30	65.5	60.5	66.7	61.7
23:00	71.8	66.8	66.9	61.9
23:30	70.7	65.7	67.8	62.8
0:00	68.4	63.4	69.4	64.4

Table C4. Nonlinearity Test Signal: Equipment Schedule for system B

	Day 1			Day 2		
Time	Stage	Lighting	Base board	Stage	Lighting	Base board heat
0:00	5	stage 1 and 2	Stage 1	3	stage 1 and 2	off
0:30	3	stage 1 and 2	off	2	stage 2	off
1:00	2	stage 2	off	5	stage 1 and 2	Stage 1
1:30	4	off	Stage 1	2	stage 2	off
2:00	3	stage 1 and 2	off	5	stage 1 and 2	Stage 1
2:30	4	off	Stage 1	4	off	Stage 1
3:00	2	stage 2	off	3	stage 1 and 2	off
3:30	6	off	Stage 2	6	off	Stage 2
4:00	1	Stage 1	off	5	stage 1 and 2	Stage 1
4:30	6	off	Stage 2	2	stage 2	off
5:00	5	stage 1 and 2	Stage 1	1	Stage 1	off
5:30	2	stage 2	off	5	stage 1 and 2	Stage 1
6:00	0	off	off	3	stage 1 and 2	off
6:30	5	off	off	6	off	Stage 2
7:00	1	Stage 1	off	3	stage 1 and 2	off
7:30	6	off	Stage 2	6	off	Stage 2
8:00	6	off	Stage 2	3	stage 1 and 2	off
8:30	5	stage 1 and 2	Stage 1	3	stage 1 and 2	off

9:00	5	stage 1 and 2	Stage 1	2	stage 2	off
9:30	2	stage 2	off	4	off	Stage 1
10:00	2	stage 2	off	2	stage 2	off
10:30	5	stage 1 and 2	Stage 1	2	stage 2	off
11:00	3	stage 1 and 2	off	4	off	Stage 1
11:30	2	stage 2	off	5	stage 1 and 2	Stage 1
12:00	5	stage 1 and 2	Stage 1	2	stage 2	off
12:30	6	off	Stage 2	2	stage 2	off
13:00	2	stage 2	off	0	off	off
13:30	2	stage 2	off	2	stage 2	off
14:00	5	stage 1 and 2	Stage 1	6	off	Stage 2
14:30	4	off	Stage 1	2	stage 2	off
15:00	1	Stage 1	off	2	stage 2	off
15:30	5	stage 1 and 2	Stage 1	4	off	Stage 1
16:00	2	stage 2	off	3	stage 1 and 2	off
16:30	5	stage 1 and 2	Stage 1	5	stage 1 and 2	Stage 1
17:00	5	stage 1 and 2	Stage 1	2	stage 2	off
17:30	2	stage 2	off	4	off	Stage 1
18:00	2	stage 2	off	3	stage 1 and 2	off
18:30	1	Stage 1	off	5	stage 1 and 2	Stage 1
19:00	3	stage 1 and 2	off	0	off	off
19:30	6	off	Stage 2	1	Stage 1	off
20:00	5	stage 1 and 2	Stage 1	6	off	Stage 2
20:30	5	stage 1 and 2	Stage 1	6	off	Stage 2
21:00	4	off	Stage 1	0	off	off
21:30	2	stage 2	off	5	stage 1 and 2	Stage 1
22:00	4	off	Stage 1	4	off	Stage 1
22:30	0	off	off	1	Stage 1	off
23:00	4	off	Stage 1	2	stage 2	off
23:30	2	stage 2	off	1	Stage 1	off
0:00	6	off	Stage 2	4	off	Stage 1

The equipment schedules will include baseboard heat and heating coil heat to represent 6 levels of internal load states. The detailed settings need to be determined with ERS stuff.

In the meantime, the building cooling energy consumption and power for these four rooms in system A will be measured at each time step. In ERS system, the cooling energy consumption for the whole system can be directly measured. The energy for each room will be calculated based on the discharge air flowrate and temperature change.

3. Response time test

In this study, the building zone temperature is chosen as the measurement in this response time experiment. Since both the HVAC system capacity and the building thermal mass could affect its zone temperature response time, two tests are performed to evaluate the response time respectively. The first one is to change the zone temperature setpoint after the building zone temperature has reached a steady state, and then measure the time between the beginning of the setpoint change and when the zone temperature reaches the 95% of the new setpoint. This response time is a combination of both building thermal mass and HVAC system capacity. The other one is to switch off the HVAC system at night, when weather disturbances are minimal, and measure the time that the zone temperature takes to decrease to a steady state (or nearly a steady state). The nearly steady state is defined as less than 0.5% of the state change in 15 minutes. This second test evaluates the building thermal mass when the weather disturbances are minimized.

For the real experiment, the temperature response time will be firstly determined through historical data examining to check how long it will take for the ERS system to reach to the new steady state after a temperature setpoint change. The temperatures at test 2 need to be based on the outside temperature during the testing days. 5, 8, and 10 degree differences between the inside and outside temperature will be maintained though cooling and heating system. If the outside temperature is low, heating system can be used to maintain the indoor temperature higher than the outdoor temperature.

Table C5. System response time test results

Setpoint change	**Temperature change time, $T_{0.95}$ (test 1)**		
	65 ºF – 77 ºF	72 ºF - 77ºF	72 ºF - 82 ºF
Time (min)	48	15	55
Setpoint change	72 ºF – 68 ºF	77 ºF - 72 ºF	82 ºF – 72 ºF
Time (min)	24	18	27
Initial temperature	**Temperature decay time, $T_{0.95}$ (test 2)**		
	70 ºF	77 ºF	85 ºF
Time (min)	~	12	93

4. System excitation signal generation

At day 4, during the normal testing, the system nonlinearity and response time will be calculated based on the experiment data. Then the system excitation signals will be generated.

5. System excitation for identification

After the system nonlinearity and system response time test finished, the system nonlinearity and response time will be calculated. And then the system excitation signal will be generated following the system identification model adaptation methods. After the calculation, a new temperature setpoint and equipment schedule excitation plan will be provided similar to Table C1 and then the building cooling energy consumption will be measured as each time step (30 minutes).

6. Normal day operating

The normal day operating: temperature setpoints, base board heat, lighting schedule will use the typical ERS settings, which need to be determined with ERS stuff.

Vita

Xiwang Li was born in Hunan, China. He received his bachelor's degree of Building Environment Engineering from Tongji University, Shanghai, China in 2010. He joined the Department of Civil, Architectural, and Environmental Engineering at Drexel University in 2010. During his Ph.D. study he focused on developing the framework to enable real-time autonomous robust and optimal building energy system operation decisions for temporally and spatially distributed building clusters. He has published multiple research works on these topics. He is also a member of many professional affiliations such as ASHRAE, IEEE, and ASME.

Selected Publications

1. Li, X., and Wen J. "Review of Building Energy Modeling for Building Control and Operation", Renewable & Sustainable Energy Reviews, 2014. 37: p. 517-537.
2. Li, X., Wen J. "Building Energy Consumption On-line Forecasting Using Physics Based System Identification", Energy and Buildings, 2014. 82: p. 1-12.
3. Li, X., Tan, H., Rackes, A., Carbon footprint analysis of student behavior for a sustainable university campus in China, Journal of Cleaner Production, Volume 106, 1 November 2015, Pages 97-108
4. Li, X., Wen, J., Bai, E.W., (2015) "Developing a Whole Building Energy Forecasting Model for On-line Operation Optimization using Proactive System Identification", (Under review, Applied Energy)
5. Li, X., Wen, J., Wu, T. "Comparison of On-line Building Energy Forecasting Model Using System Identification Method and Other Inverse Modeling Methods", in ASHRAE 2015 Annual Conference; Jun. 27 –Jul. 1, 2015, Atlanta, GA, USA.
6. Li, X., Wen, J., Bai, E.W. "Building Energy Forecasting Using System Identification based on System Characteristics Test", in 2015 Workshop on Modeling and Simulation of Cyber-Physical Energy Systems; Apr. 13-17, 2015,Seattle, WA, USA.
7. Li, X., Wen, J., Wu, T. "Net-Zero Energy Impact Building Clusters Emulator for Operation Strategies Assessment". ASHRAE 2014 Annual conference; Jun. 28 –Jul. 2, 2014, Seattle, WA, USA.
8. Li, X., Wen, J. "Building Energy Consumption On-line Forecasting Using System Identification and Data Fusion". ASME 2014 Dynamic Systems and Control (DSC) Conference, Oct. 22-24, 2014, San Antonio, Texas, USA.